TV 視聴実データを用いた番組面白さ関数の提案

上田光流 [1], 川畑泰子 [2], 石井　晃 [1], 榎屋裕三 [3]
[1] 鳥取大工, [2] 東大情報理工, [3] ビデオリサーチ

TV 番組における瞬間視聴率とは何を意味するだろうか．一般の芸能誌／紙の報道では，瞬間最高視聴率の瞬間とはその番組で一番面白かった瞬間という意味で使っている．しかし，ある時刻 t で瞬間最高視聴率ということはその時刻 t での場面をきっかけに視聴者がその番組から離れて行ったという意味にもなる．とは言え，まったく面白くない場面が続いていたら，その前からずっと視聴率は落ち続けるので，その時刻 t が瞬間最高視聴率となることもない．ということで，『瞬間最高視聴率の意味』とは，単純ではないと思えてくる．

先行研究として，瞬間視聴率のデータがあると仮定しての物理学の立場からの数理モデルが喜屋武寛之，井上純一によって提出されている[1, 2]．ただ，喜屋武・井上モデルでは局から局へのチャンネル変更を熱統計力学のギブス分布になぞらえて定義していて，この仮定の妥当性には議論があるところであろう．また，喜屋武・井上モデルでは番組自体の面白さよりも TV 番組に挟み込まれる広告の適切な時間配分の決定に注目していて，本稿で考える，瞬間視聴率データから番組の面白さの時間的変化を探る事とは，目的が異なっている．

本稿では，上記の先行研究とはまったく別に，簡単な数理モデルで瞬間視聴率を考えてみよう．そこから，瞬間視聴率と時々刻々の番組の面白さの関係を数理で裏付けることを目指す．まず，ある時間Δ t の間での視聴者数の変化とは，次のような意味であることに異論はないだろう．

視聴者数の変化　＝　新たに視る人の数　ー　視るのを止めた人の数

この人数の変動は瞬間瞬間で変わっていくので，それぞれが時間の関数であると考えて

視聴者数の変化(t)　＝　新たに視る人の数(t)　ー　視るのを止めた人の数(t)

とするべきであろう．ここで記号として

視聴者数の変化(t) = $\Delta V(t)$

新たに視る人の数(t) = $Nin(t)$

視るのを止めた人の数(t) = $Nout(t)$

と書くと，上の式は

$$\Delta V(t) = [Nin(t) - Nout(t)]\Delta t \qquad (1)$$

となる．

さて，次にそれぞれの中身を考えてみよう．時刻 t から新たに視る人の内訳は，時刻 t で初めて TV を点けた人と，他局の番組から移ってきた人であろう．

新たに視る人(t) = TV を見始めた人(t) + 他局から移ってきた人(t)

同様に，時刻 t でこの番組を見るのを止めた人は，この時刻で TV 自体をオフにした人と，他局に移った人であるはずなので,

視るのをやめた人(t) = TV を消した人(t) + 他局に移った人(t)

となる．

記号を導入していこう．

時刻 t で TV を見始めた人の数(t) = $N_{on}(t)$

時刻 t で TV を消した人の数(t) = $N_{off}(t)$

時刻 t で他局 j から移ってきた人の数(t) = $N_{ji}(t)$

時刻 t で他局に移った人の数(t) = $N_{ij}(t)$

ここで現在注目している番組の局を i としている．すると,

$$N_{in}(t) = N_{on}(t) + N_{ji}(t) \qquad (2)$$

$$N_{out}(t) = N_{off}(t) + N_{ij}(t) \qquad (3)$$

となる．

TV 局の数は国によって異なる．例えばベトナムは 1 局だけであるし，メキシコ，ミャンマーや北朝鮮のように TV 局が 2 局しかない国では上記の式で十分だが，日本の場合は多くの地方で NHK2 局と民放が 5 局，他に BS もあるので，丁寧に数えると 15 局くらいあることになる．従って，上の式での”他局 j”は 1 つではない．また, TV を点けた時に最初に出るチャンネルは，視聴者の意図どおりではなく，前回に消した時のチャンネルである可能性が高い．そこで,

局の数 ＝ B

として，確率が平等とすると, TV を点けて局 i が出現する数は

$$N_{on}(i,t) = \frac{1}{B} N_{on}(t) \tag{4}$$

となるだろう．チャンネルの出現確率を平等とするのは乱暴な近似ではあるが，この問題はあとで考察する．点ける場合と逆に TV を消す人の数 $N_{off}(t)$は今視ているチャンネルが面白くなくて（あるいは次に予定があって）TV を消すので, N_{off}そのままである．

他局 j 全部について考えると，時刻 t からこの番組を見始める人の数は

$$N_{in}(t) = \frac{1}{B} N_{on}(t) + \sum_j N_{ji}(t) \tag{5}$$

となる．和は他局全部について取る．一方，時刻 t でこの番組を視るのをやめる人の数は

$$N_{out}(t) = N_{off}(t) + \sum_j N_{ij}(t) \tag{6}$$

従って，時刻 t で局 i での番組を視る人の数の変化分は

$$\Delta V_i(t) = \left[\frac{1}{B} N_{on}(t) + \sum_j N_{ji}(t) - N_{off}(t) - \sum_j N_{ij}(t)\right] \Delta t \tag{7}$$

となる．これと同様の方程式が TV 放送局の数だけ立てられて，それの連立が瞬間視聴率の数理モデルとなる．もちろん，このままでは単なる枠組みであって，まだ中身は何も入っていない．

さて，次にどのような場合にチャンネルを変えるか，つまり番組を視るのをやめるかを考えてみよう．まず，実際のビデオリサーチ社のデータは，個々の視聴者ごとにどのチャンネルからどのチャンネルに移ったかが記録されている．この場合は，以下の量は直接測定されていると考えていい．この場合には，テレビのスイッチを入れた瞬間にどの局を視ていたか，という不確定要素が大きい問題は考えなくていいことになる．

時刻 t で他局 j から局 i に移ってきた人の数(t) = $N_{ji}(t)$

時刻 t で局 i から他局に移った人の数(t) = $N_{ij}(t)$

チャンネルを変える原因として，番組が面白くないからというのと，裏番組が面白そうだからという２つの理由があるだろう．まず，面白さから考える．番組の面白さを１００％満喫している状態を１と考えよう．そして，時刻 t での i 局の番組の面白さが測れるとして，それを Am(i, t)で表そう．

番組の面白さ(t) = Am(i, t)

すると，番組のつまらなさは Am(i, t)が負の値となることで表せるだろう．

さて，視ている番組から視聴者が離れてしまう場合，番組がつまらないからと考えるのが素直である．とすると，他局に移る人の数は，番組のつまらなさに比例すると考えられる．面白さの上限を１と仮定しているので，これは単純にこの番組からチャンネルを移そうとする確率である．すると，今, i 局の番組を視ている人の数を$V_i(t)$として，他局に移る人の数は次のように表せるであろう．

$N_{ij}(t) = -Am(i,t)c_jV_i(t)$ (8)

これから，時刻 t で i 局から他局へ移った人の総数は

$\Delta N_i(t) = \sum_j N_{ij}(t) = -\sum_j Am(i,t)c_jV_i(t)$ (9)

となる．

逆に，時刻 t で他局 j から移ってくる人はこれの裏返しであると考えられる．つまり，

$N_{ji}(t) = -Am(j,t)c_iV_j(t)$ (10)

これから，時刻 t で他局から i 局に来た人の数は

$\Delta N_i(t) = \sum_j N_{ji}(t) = -\sum_j Am(j,t)c_iV_j(t)$ (11)

従って, i 局を見ている人の数の変化は，以下のようになる．

$\Delta V_i(t) = \left[\frac{1}{B}N_{on}(t) + \Delta N_i - N_{off}(t)\right]\Delta t$ (12)

そして，

$\Delta N_i(t) = -\sum_j Am(j,t)c_iV_j(t) + \sum_j Am(i,t)c_jV_i(t)$ (13)

となる．ここで$\Delta N_i(t)$と$V_i(t)$がビデオリサーチが持っている既知の量として，ここから$Am(i,t)$を求めることになる．求めるには全局の視聴情報から連立で求める．

$\Delta N_i(t) = -c_i\sum_j Am(j,t)V_j(t) + Am(i,t)V_i(t)\sum_j c_j$ (14)

ここで係数c_jは他局の中から j 局を選んで移る確率である．この係数c_jは番組の事前評判やネット上での期待，そして裏番組を同時進行で見ている人からのツィートに影響される場合もあるであろう．これは他局全部で規格化されている必要がある．しかし，我々の調べでは番組放送時に同時に呟かれるツィートの影響はあまり無いというのがデータで出ている．各局（の当該時間の番組）への期待度をE_jとすると

$c_j(t) = \frac{E_j(t)}{\sum_{k\neq j}E_k(t)}$ (15)

というように，規格化された形で定義される時間の関数だろう．各番組への期待度は

期待度＝事前情報からの期待度

なる．事前情報からの期待度は，さらに番組宣伝やそれに対する視聴者の反応（ブログや Twitter など）から構成されると思われる．

仮に番組が始まる時間を時刻 t =0 とし，期待の盛り上がりを$I(t)$という関数で表せるとすると（これはヒット現象の数理モデルで扱う量と似ている），

事前情報からの事前の盛り上がり　=　$\int_{-\infty}^{0} I(t)\mathrm{dt}$

これらは原理的には視聴者一人一人，異なる係数がかかって，期待度に結びつく．

$$E_j(t) = \alpha_j \int_{-\infty}^{0} I_j(t)\,dt + \beta_j \tag{16}$$

ここで第一項が事前のツィート等で見られる番組への盛り上がり．第二項はツィートでは測れない，番組に対する期待を表す．

つまり，ネット等による各番組の期待確率$c_j(t)$は

$$c_j(t) = \frac{E_j(t)}{\sum_{k\neq j} E_k(t)} \tag{17}$$

$$E_j(t) = \alpha_j \int_{-\infty}^{0} I_j(t)\,dt \; + \; \beta_j \tag{18}$$

で決められる．ここで重要となる，

まとめると，求めるべきは

$$\Delta N_i(t) = -c_i \sum_j Am(j,t)V_j(t) + Am(i,t)V_i(t) \sum_j c_j \tag{19}$$

であり，ここで各番組の期待確率$c_j(t)$は

$$c_j(t) = \frac{E_j(t)}{\sum_{k\neq j} E_k(t)} \tag{20}$$

$$E_j(t) = \alpha_j \int_{-\infty}^{0} I_j(t)\,dt \; + \; \beta_j \tag{21}$$

で求めて，これを上の式(19)に代入する．

３　分析方法

3.1　計算方法

ビデオリサーチから提供のデータは関東の 600 世帯についてのテレビ放送に関する視聴数データで，個別集計している NHK 総合,E テレ，日本テレビ，TBS，フジテレビ，テレビ朝日，テレビ東京の 7 局のデータである．視聴数が極端に少ない E テレを除くと 6 局となる．この 6 局を独立として連立させ，(19)式を解くべきであるが，現段階では(8)式，(10)式で必要な，A 局から B 局へ移動する視聴者数のデータ自体ではなく，１局と，それ以外の 5 局の合算という 2 局で計算を行った．目的は視聴者数自体の増減と，ここで求める面白さの関数 Am(i,t)が一致するかどうかであり，一致しなければ面白さの関数を研究する意味があることになる．

今回は，２局による数理モデルを考えていくので，各局（の当該時間の番組）への期待

度をE_jとすると

$$c_j(t) = \frac{E_j(t)}{\sum_{k \neq j} E_k(t)}$$

というように，規格化された形で定義される時間の関数は，2 局しかない場合，条件は次のようになる．

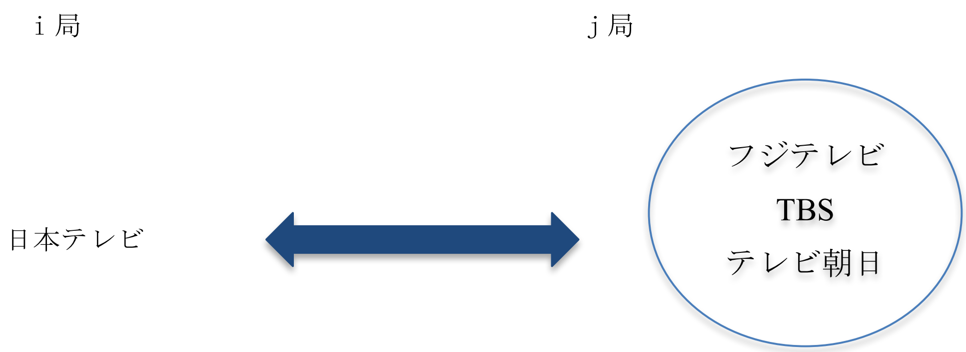
$$C_i(t) = 1 \qquad C_j(t) = 1 \tag{22}$$

式（22）により，視聴率数理モデルの式はより簡単化された．エクセルを使用して，連立方程式を解き分毎の面白さの関数を出した．

3.2　分析方法

セクション 3.1 でも述べたように，2 局による数理モデルを考える．基準の 1 局に対し，2 局，3 局を考えるとなると視聴率数理モデルの式が複雑になると考え，まずはC_i，C_jが 1 として考えることができる 2 局による数理モデルでやることにした．また，視聴率のデータはビデオリサーチ社から頂いたものを使用した．

例）



3.3　計算結果

図にある人気ドラマの場合の計算を示す．これは当該放送局以外の 5 局を「もう 1 局」とした場合の，ドラマ番組自体についての計算である．赤い線が従来から視聴率の算定基準になっている，視聴者数であり，番組の 1 時間を通じて増減していることがわかる．これは 1 分ごとの値である．(19)式で計算された面白さの関数$Am(i,t)$が青線である．図を見てわかるように，面白さの関数はかなり激しく上下している．上に伸びたピークが面白いポイント，下に伸びたピークが面白くないポイントである．全てのピークで説明がつくわけではないが，面白いポイントで視聴者数の変化が上向きとなり，面白くないポイントで逆に視聴者数が下向きになっている箇所が見て取れる．

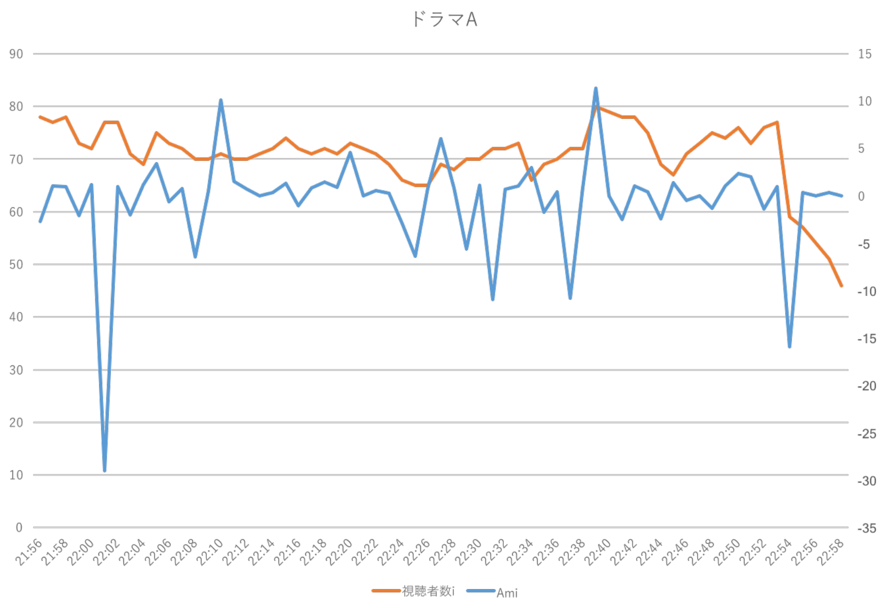


図１　番組視聴者数と面白さ関数．赤線が番組視聴者数で縦軸は左側．青線が番組面白さ関数で縦軸は右側．番組は人気ドラマの関東地区放送．

４　考察

図１で示されるように，番組面白さ関数は視聴者数の増減と必ずしも一致しない．その理由として，視聴者数は毎分の流入数と流出数の差で決まるが，(19)式の計算は流入と流出を分けて考えていることが挙げられる．つまり，視聴者数だけでは，例えば２０人流入し，同時に２０人流出した場合，変化無しとなってしまうが，本研究の理論だと流入と流出を別々に扱うので，流出はその番組の１分前が面白くなかったことに，流入は１分前に他局の番組が面白くなかったことに対応する．実データを見ると１分ごとに流入，流出はかなり激しく増減していて，このため，面白さ関数は視聴者数とはかならずしも一致する動きにならないと考えられる．

しかし，今回我々は１７番組について，６局それぞれで（２局の近似で）計算したが，必ずしも図１に示すほど視聴者数の変化と面白さ関数の変化が一致しないケースも見られる．少なくとも視聴者数のピークと面白さ関数のピークが一致しないことは確認できたが，実際は面白さピークと面白くないピークが具体的に番組のどの場面であったかの確認も必要となる．また，６局相互で移動する視聴者数のデータを得て，６局独立で連立させて(19)式を解くことで，より精度の高い番組面白さ関数の計算ができるものと考えられる．

また，各局の面白さの関数の関係性分析では，ある局が面白い時にある局がつまらなくなっていることがある．これからこれらの局で流入，流出が行われたのではないかと考え

られる．どの局からどの局に流入，流出があったかがわかると何かに生かせるかもしれない.

５　結論

視聴率の根拠となっている番組の１分ごとに視聴者数の増減はかならずしも面白さのポイントに対応していないという動機から，視聴行動の簡単な数理モデルを提案した．１局と他局全てという２局近似で，番組面白さの関数を計算し，これが視聴者数の増減と一致するものではないことが示された．しかし番組面白さの関数を実用化するには６局を独立に連立させた計算が必要であり，今後は 6 局相互の間での流入・流出データを使用してより精緻な面白さ関数を計算していく.

参考文献

[1] 喜屋武寛之，井上純一，「TV コマーシャル市場の確率モデル」電子情報通信学会技術研究報告. NLP,　非線形問題 112(389),　61-66,　2013-01-17

[2] 喜屋武寛之　「テレビコマーシャル市場の確率モデル」北海道大学情報学研究科複合情報学専攻 複雑系工学講座 混沌系工学研究室修士論文 2013

[3] A.Ishii,　H.Arakaki,　N.Matsuda,　S.Umemura,　T.Urushidani,　N.Yamagata and N.Yoshida; The 'hit' phenomenon: a mathematical model of human dynamics interctions as s stochastic process,　New Journal of Physics 14 (2012) 063018 (22pp)

[4] 吉田就彦，石井晃，新垣久史「大ヒットの方程式」ディスカヴァートゥウェンティワン 2010.