２項分布分析チャートを活用した高等学校統計教材の開発

Development of Teaching Materials on Statistics
at Upper Secondary School Level Using Binomial Distribution Analysis Chart

成　田　雅　博 *
NARITA Masahiro

要約：本研究では，統計教材でとりあげる統計モデルとして２項分布モデルを採用する際，実体的イメージの形成をうながす導入教材としての「末広がりスゴロク」を解説した。次に，多様な現象の説明から２項分布のパラメーターの抽出を支援する枠組みとして開発された２項分布分析チャートを解説し，２項分布にしたがう現象の分類を試みた。

キーワード：数学，統計，データの分析，２項分布，２項分布分析チャート

１．はじめに

新学習指導要領においては，教科数学に統計に関する教育内容が多く盛り込まれている。また統計情報に係るリテラシー教育の観点からは，高等学校教科情報や，各教科の中で扱うべき内容としての情報教育にも統計に関する教育内容が含まれていると考えられる。具体的な学習活動としては，探索的データ解析（ＥＤＡ　Exploratory Data Analysis）の諸手法や品質管理に利用されるツールを使って，PDCA サイクル（Plan（計画）－Do（実施・実行）－Check（点検・評価）－Act（処置・改善））による統計的問題解決過程を通して，このような手法の意義の理解，技能の修得に重点がおかれることになるが，マスコミや各種調査研究の報告書でふれられることの多い推測統計（信頼区間や統計的検定）の結果を批判的に受容・評価する活動も行うことになると考える。このような推測統計の教材としては，正規分布をはじめとする連続分布にしたがう事象を題材とした教材が多く取り上げられているが，連続分布は離散分布にくらべ，標本統計量の確率に関する認識が困難であり，高等学校や大学教養課程において推測統計を教える際の障害になっていると思われる。そこで，本研究では，より認識が容易と考えられる２項分布にしたがう現象に焦点をあて，教材研究を支援する資料を整理することとする。

２項分布にしたがう現象の教材としての価値としては，以下をあげることができる。

（１）離散分布であるため，正規分布等の連続分布と異なり，確率変数の値がちょうどぴったりの値の確率が存在すること。

（２）確率値を加法法則・乗法法則・組み合わせ論的知識だけから計算できること。

（３）ある理想的な条件のもとで多くの現象が２項分布として扱えること。

具体的にはまず，単一試行の結果である事象が２つしかないベルヌーイ試行を独立に繰り返したと解釈される２項分布現象の導入に適切と思われる教材「末広がりスゴロク」と，現象から２項分布モデルのパラメーターを抽出することを容易にする２項分布分析チャートを説明する。次に，さまざまな２項分布モデルの使える問題状況を分類化して提示する。

* 附属教育実践総合センター

２．２項分布導入教材としての「末広がりスゴロク」

２項分布は，単一のベルヌーイ試行が独立に繰り返し行われた際の「繰り返し試行」における「成功」回数の確率分布として定義される。この２項分布の実体的イメージ（高村 1987）の形成をうながす典型的な活動を行う教材として，以下の「末広がりスゴロク」が開発されている（成田 2007）。

２項分布モデルを記述するパラメーターは，p：単一試行における「成功」の確率，n：繰り返し回数の２つであり，そのモデルは B (n, p) と表記されることが多い。また，「繰り返し試行」における「成功」回数が２項分布における確率変数 X である。この教材「末広がりスゴロク」における２項分布モデルは B (6, 1/6) に従い，確率変数 X は，６回の試行のうち１の目が出る回数である。また，特定の確率変数値 χ の確率値は

$$p(x) = {}_6C_x \times \left(\frac{1}{6}\right)^x \times \left(\frac{5}{6}\right)^{6-x}$$

であり，Excel の関数では，=BINOMDIST（x,6,1/6,FALSE）により値を計算することができる。

【末広がりスゴロク】

・使うもの・・・・「末広がりスゴロク・ボード」，サイコロ１個，チップ２０枚程度
・ゲームの手順
（１）ディーラーは，「あ」にチップをおき，サイコロを６回つづけてふり，出た目に応じてチップをひとマスずつすすめます。１の目が出たら ↘ に，１以外（２～６）の目が出たら ↙ にすすめます。
（２）プレイヤーは，ディーラーがサイコロをふりはじめる前に，チップがどこに到達するかを予想してそこにプレイヤーのチップをおきます。
（３）プレイヤーがディーラーのチップの到達先を当てることができたら，プレイヤーがチップを２枚ともとり，はずれたらディーラーが２枚のチップをとります。ただし，６回ふり終わる前にディーラーのチップが通過しただけの場合は「当てた」とはみなしません。

ここまでを１セットとして，結果を表に記録します。

セット	1	2	3	4	5	6	7	8	9	10
チップ到達予想										
実際のチップ到達箇所										

プレイヤーが当てた回数	

☆　質問　☆
【末広がりスゴロク】で，一番到達しやすいところはどこだと思いますか。
また，そこには，10 セットやってみると平均して何回くらい到達すると思いますか。

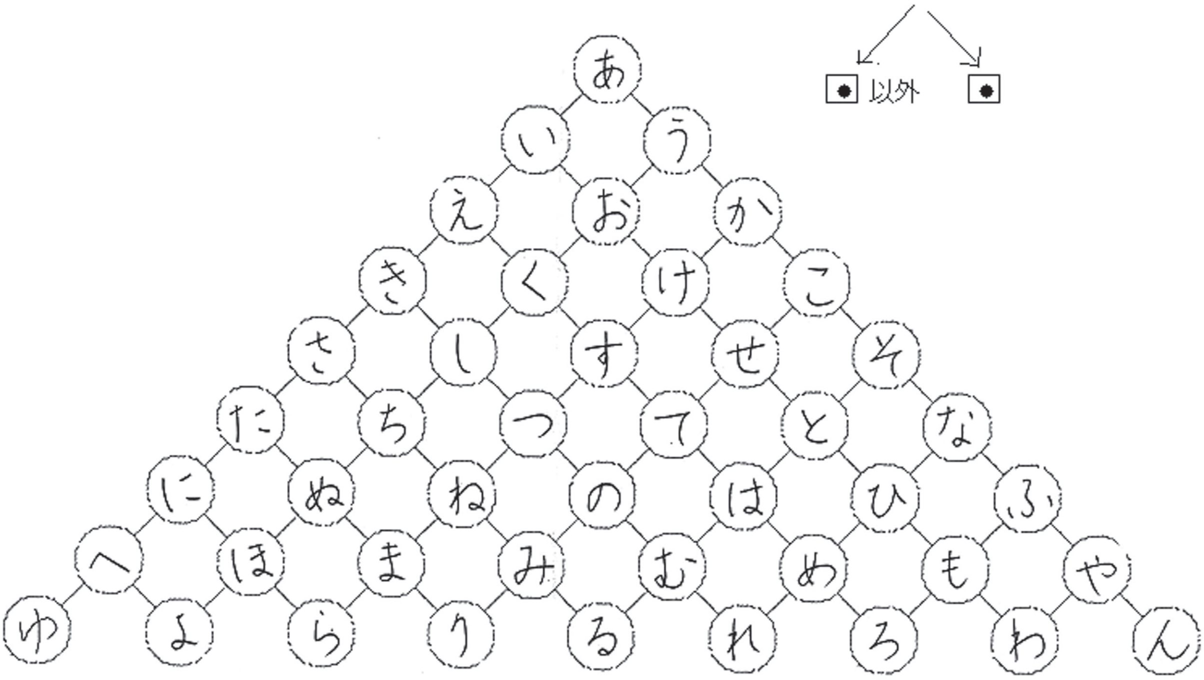

図１　末広がりスゴロク・ボード

この教材は，パチンコ台のような玉が左右に 1/2 の確率で落ちていき，玉の落ちる分布が左右対称になるクインカンクス，またはゴルトンボード（Pierce : 2011）とは異なり，左右非対称であり，より典型的な現象であると考えられる。また，学習者の活動が，単一試行であるサイコロ１回投げと，その単一試行を６回繰り返した１セットを試行とみなし最終的にどこに到達するかに注目する繰り返し試行との「二重性」を自然と意識することができる教材として構成されている。

右下に１つ進む１の目がでる確率が 1/6 であり，その試行を独立に６回繰り返すので，平均的には１回右下にすすみ，残りの５回が左下にすすむ「ぬ」にもっとも多く到達することが予想される。この末広がりスゴロクの到達確率を計算してみると，表１のように予想どおり「ぬ」の確率が大きいが，「に」のように，６回のうち６回とも１以外の目が出る確率も多く，いつもプレイヤーが「ぬ」と予想した場合約 60% の確率で予想が外れることになり，ゲームとしてはある程度の意外性が感じられるものと考えられる。

表１　末広がりスゴロクの１セットで到達する確率

ボードの到達場所	１の目の出る回数 X	１セットでそこにチップが到達する確率	10 セットのうち４回以上そこにチップが到達する確率
に	X=0	0.3349	0.4450
ぬ	X=1	0.4019	0.6224
ね	X=2	0.2009	0.1225
の	X=3	0.0536	0.0013
は	X=4	0.0080	0.0000
ひ	X=5	0.0006	0.0000
ふ	X=6	0.0000	0.0000

３．２項分布モデルに関するパラメーター抽出を支援する２項分布分析チャート

２項分布分析チャートは，「単一試行」における事象・確率と，「繰り返し試行」における事象・確率とを重ねて整理したものである。たとえば，導入教材「末広がりスゴロク」をこのチャートで整理すると以下のようになり，パラメーター $p = 1/6$, $n = 6$ を抽出・整理することができる。

＜１＞【末広がりスゴロク】（サイコロを６回ふる）B(6 , 1/6)

	試行	事象	確率
単一試行	サイコロを１回ふる	１の目が出る (↘ へすすむ)	p（１の目） = p(↘)= $\frac{1}{6}$

↓ ＜独立にくりかえし＞

「６回くりかえし」の試行	サイコロを６回ふる	１の目が x 回出る (↘へ x 回すすむ)	$p(x) = {}_6C_x \times \left(\frac{1}{6}\right)^x \times \left(\frac{5}{6}\right)^{6-x}$ =BINOMDIST(x ,6,1/6,FALSE)

次の現象に対して，２項分布分析チャートを適用してみよう。ただし，各打席における打率は不変であり，各打席は独立であるという「理想化」された仮定のもと，はじめて２項分布モデルが適用でき，確率が計算できる。

＜２＞【野球の試合のヒット数】

・ある選手の打率（ヒットを打つ割合）が 0.3 であると仮定します。
この選手がある試合で５回打席にたったとき，ちょうど２本ヒットを打つ確率はいくらですか。

＜２＞【野球の試合のヒット数】B (5 , 0.3)

	試行	事象	確率
単一試行	１つの打席に立つ	ヒットを打つ	p（ヒットを打つ）＝ 0.3

↓ ＜独立にくりかえし＞

「５回くりかえし」の試行	打席に５回立つ	ヒットを x 回打つ	$p(x) = {}_5C_x \times (0.3)^x \times (0.7)^{5-x}$ = BINOMDIST(x ,5,0.3,FALSE)

＜３＞【飛行機予約でのオーバーブッキング】

定員 261 人の飛行機に，キャンセルを見込んで 265 人の予約を受けた。
キャンセル率２％のとき，乗れない人の出る確率を求めてください。

この場合も，乗客全員が個人旅行でありそれぞれ独立にキャンセルを決め，キャンセルの確率も全員同じである，という非現実的な仮定のもと，はじめて２項分布モデルを適用することができ，確率を計算することができる。

＜３＞【オーバーブッキング】 (265 , 0.02)

	試行	事象	確率
単一試行	予約を受けた１人がキャンセルするかどうかを調べる	その１人がキャンセルした	p（キャンセル） = 0.02

↓ ＜独立にくりかえし＞

「２６５回くりかえし」の試行	２６５人についてキャンセルかどうか調べることをくりかえす	キャンセルが x 人である	$p(x)={}_{265}C_x \times (0.02)^x \times (0.98)^{265-x}$ =BINOMDIST(x ,265,0.02,TRUE)

４．２項分布現象の分類

上記のような多様な２項分布モデルで解釈可能な現象に対して，現在暫定的に，以下のようにＡ１～Ａ５およびＢ１～Ｂ３の仮説的な分類を設けている。一般に，Ａ１からＡ５にすすむにしたがい，また，Ｂ１からＢ３にしたがい，理解が困難になることが予想される。

（A）単一試行問題の文脈

Ａ１　ギャンブル的・同様に確からしい
・・・コイン・サイコロ・トランプ

Ａ２　一般的・同様に確からしい
・・・実力の同じチームの対戦・誕生月

Ａ３　事象の確率値，または事象の相対度数の明示的提示
・・・命中率・不良品の比率・成功率

Ａ４　一定の時間帯における事象の確率値，または事象の相対度数の明示的提示
・・・ある期間における故障率・生存率 / 死亡率・事故発生率・火災発生率

Ａ５　空間分布・時間分布への再構成
・・・特定の粒子が特定の長方形に入る

（B）くりかえし試行における試行間の関係

Ｂ１　継時的
・・・１個のサイコロを２回ふる

Ｂ２　同時的
・・・２個のサイコロを１回ふる

Ｂ３　空間的・時間的分解操作
・・・空間ブロックまたは時間ブロックへの分解を含む

この２つの観点で２項分布現象を分類した表を以下にあげる。

表２　試行的な２項分布現象の分類

現象の概要	分類Ａ	分類Ｂ	B (n, p)
末広がりスゴロク：サイコロを６回ふる。１の目の１回だけ出る確率は？	Ａ１	Ｂ１	B (6, 1/6)
「2 項分布説明器」・クインカンクスまたはゴルトンボード。左右に分ける杭が７段階ある場合	Ａ１	Ｂ１	B (7, 1/2)
10 個のサイコロを一度にふる。１の目が３個だけ出る確率は？	Ａ１	Ｂ２	B (10, 1/6)
あるクラスにいる 30 人のうち３月生まれの人が３人いる確率は？	Ａ２	Ｂ２	B (30, 31/365. 25)
お年玉付き年賀葉書の４等（お年玉切手シート 100 枚に２枚当たり）が 10 枚届いた。１枚当たる確率は？	Ａ２	Ｂ２	B (10, 0. 02)
実力の同じチームが７連戦する。Ａチームが４勝３敗で勝つ確率は？	Ａ２	Ｂ１	B (7, 1/2)
Ａチームが試合に１試合に勝つ確率は 0. 6。年間 130 試合中 70 勝以上となる確率は？	Ａ３	Ｂ１	B (130, 1/2)
ある選手の打率が 0. 3 のとき，ある試合で５打席で２本ヒットを打つ確率は？	Ａ３	Ｂ１	B (5, 0. 3)
アンケートの回収率が 60% のとき，400 通送って 260 通以上回収できる確率は？	Ａ３	Ｂ２	B (400, 0. 6)
献血した人が 200 人いたとき，Ｂ型が 30 人以上である確率は？	Ａ３	Ｂ２	B (200, 0. 2)
不良率２％の製品の山から 10 個とったとき，不良品が含まれる確率は？	Ａ３	Ｂ２	B (10, 0. 02)
男子の出生率が 0. 5 であるとき，４人の子どもが４人とも男子である確率は？	Ａ３	Ｂ１	B (4, 1/2)
４択問題から１つだけ正解を選ぶ問題が 10 問あったときでたらめに答えて５問正解する確率は？	Ａ２	Ｂ１	B (10, 1/4)
超能力がないと仮定したとき，100 回コインを投げて 60 回以上表裏を当てる確率は？	Ａ３	Ｂ１	B (100, 1/2)
オーバーブッキングの問題：定員 47 人の飛行機でキャンセルを見込んで 50 人の予約を受けた際，キャンセル率が５％のとき，乗れない人が出る確率は？	Ａ４	Ｂ３	B (100, 1/2)
18 歳の男性の 50 年後生存確率が 0. 78 のとき，25 人の男性のうち 50 年後 20 人以上の生存確率は？	Ａ４	Ｂ３	B (25, 0. 78)
レーズンパンのスライスの中のレーズンの個数：100 個のレーズンを入れて作ったパンを 10 枚にスライスしたときのスライスあたりのレーズン数が３個以下になってしまう確率は？	Ａ５	Ｂ３	B (100, 1/10)

５．今後の課題

本稿では，２項分布でモデル化できる現象について，チャートでの整理をしながら，仮説的に分類を試みた。今後は，２項分布に関する学習経験・能力の多様な高校生や大学生に対する上記分類カテゴリーに属する問題の正誤データをもとに，IRT（項目反応理論）等を使い，上記分類による難易度の差について検討することで，カリキュラム開発の際の知見を得たい。

付記　本研究は，平成 22-24 年度文部科学省科学研究費補助金・基盤研究（C）『２項分布にしたがう現象のモデル化を題材とする「情報の科学」カリキュラムの開発』（課題番号：22500806，代表：成田雅博）の支援を受けた。

参考文献

成田雅博 (2007)．高等学校及び大学教養課程における２項分布にしたがう現象を題材とした統計的仮説検定に関する教材開発．第３回統計教育の方法論ワークショップ．全 17 頁．(於同志社大学)

NARITA, Masahiro(2007). Teaching Materials Using Board Game and Classifying Table for Helping Understand Binomial Distribution. ISI 2007 Book of Abstracts-International Statistical Institute 56th Biennial Session（at the Lisbon Convention Centre, Lisbon, Portugal）. P.424

成田雅博 (2011)．２項分布分析チャートと２項分布に関する問題の分類．数学教育学会誌 2011 年度臨時増刊号. pp. 117-119

Pierce, Rod (2011). Quincunx Explained-Math Is Fun. http://www. mathsisfun.com/data/quincunx-explained. html.（2011年10月１日閲覧）

高村泰雄（1987）．物理教授法の研究．北海道大学図書刊行会. pp. 44-47.