

統計数理研究所 2014 年度子ども見学デー

BB 弾サンプリング実験 解説

2014 年 10 月 18 日（土）、統計数理研究所では「子ども見学デー2014¹」を開催いたしました。当日は天候にも恵まれ、多くの方々に当研究所へお越しいただきました。お立ち寄りいただいた皆様、どうもありがとうございました。

子ども見学デーでは、「じゃんけんゲーム」と「BB 弾サンプリング実験」の 2 つの催し物をご用意いたしました。ここでは、BB 弾サンプリング実験の目的と実際の実験結果について解説していきます。



BB 弾サンプリング実験
から何がわかるかな？



トースター博士



スタッツ君

目次

1. BB 弾サンプリング実験の目的.....	2
1 - 1. 全員を調べなくても良いの？ ～社会調査とサンプリングの関係～	2
1 - 2. どうやってサンプリングしているの？ ～サンプリング法いろいろ～ ...	3
2. BB 弾サンプリング実験の結果.....	6
2 - 1. 母集団と標本の関係は？ ～標本の割合とその分布～	6
2 - 2. 誤差を減らすには？ ～標本誤差、標準誤差とサンプルサイズ～	11
3. おわりに	14

¹ 当日の様子は、<http://www.ism.ac.jp/events/kodomo2014/index.html> をご覧ください。

1. BB 弾サンプリング実験の目的

さて、今回の BB 弾サンプリング実験は、一体どのような目的でおこなわれたのでしょうか。統計数理研究所のマスコット・キャラクター、トースター博士&スタッツ君と一緒に、まずは実験の目的について学ぶことにしましょう!

1-1. 全員を調べなくても良いの? ~社会調査とサンプリングの関係~



スタッツ：ねえトースター博士、昨日テレビで「日本人の国民性調査²」っていう、社会調査³についてのニュースを見たんだ。



トースター：うむ、「日本人の国民性調査」とは、統計数理研究所が 1953 年から 5 年ごとに実施している社会調査のことじゃな。たしか 2013 年に、第 13 次全国調査がおこなわれたばかりのはずじゃ。



ス：うん。その第 13 次調査の結果によると、日本人を「勤勉」、「礼儀正しい」、「親切」と考える人の割合が、過去最高になったんだって。でも、これって何か変じゃない?



ト：はて、どうしてスタッツ君はそう思うんじゃ?



ス：だって、僕は調査を受けていないよ。友達にも聞いてみたけれど、だれもそんな調査には答えていないって。だから気になって、インターネットでも調べてみたんだ。そうしたら今回の調査では、3170 人しか調査に回答していないんだって。日本には 1 億人以上の人がいるんだよ。それなのに 3170 人を調べただけで、日本で「勤勉」、「礼儀正しい」、「親切」だと思う人が増えたなんて、本当にいえるのかなあ?



ト：ほほう、それは良い質問じゃ。実は「日本人の国民性調査」のように、統計的（定量的）な分析⁴を念頭においておこなわれる社会調査には、大きく

² 「日本人の国民性調査」の詳細な情報は、<http://www.ism.ac.jp/kokuminsei/> をご覧ください。なお、第 13 次全国調査の結果が公表されたのは、2014 年 10 月 30 日なので、実際には「子ども見学デー」よりも後になります。

³ 各種の社会調査については、ミネルヴァ書房刊『世論をさがし求めて—陶片追放から選挙予測まで—（西平重喜 2009）』にわかりやすくまとめられています。また、今回の BB 弾サンプリング実験に関連する事柄も述べられています。

⁴ 他に個々の事例を丹念に調べること（= 定性的な分析）による、「個性記述的」な調査も

分けて2つの方法があるんじゃないよ。1つ目は、関心を持っている対象(=「母集団」)を全て調べる「全数調査(悉皆調査)」じゃ。国がおこなう「国勢調査」は、代表的な全数調査じゃな。全数調査をすれば、実態を正確に把握できるんじゃないが、そのためには莫大な費用と時間と人員が必要になってしまう。だから、実際に大規模な全数調査がおこなわれることは、それほど多くはないんじゃないよ。



ス：へえ、ということは「日本人の国民性調査」は、全数調査ではない方法でおこなわれていたということなのかな？ それじゃあ博士、もう1つの方法って一体どんなものなの？



ト：うむ。スタッツ君のいうとおり「日本人の国民性調査」は、母集団から取り出された一部分(=「標本」⁶)を調べることで、全体のおおよその状態を推測する「標本調査」という手法でおこなわれておる⁷。なお、こうした調査のために母集団から標本を取り出すことを、「サンプリング(標本抽出)」というんじゃないよ。

1-2. どうやってサンプリングしているの？ ～サンプリング法いろいろ～



ス：う～ん、「日本人の国民性調査」が標本調査であることは良いけれど、どうやって標本となる人を選んでいるの？ ある特定の意見を持っている人ばかりが標本になっていたら、やっぱり日本全体の意見とは食い違いが出てくるよね？ なんだか納得いかないなあ。



ト：そうじゃな。だから、標本の選び方(=「標本抽出法」⁸、またはサンプリング法)はとても重要なんじゃないよ。そして、この標本抽出法にも大きく分けて2つの方法があるんじゃないよ。1つ目の方法は、調査者が自らの理論や主観に

あります。これについては、新曜社刊『フィールドワーク一書を持って街へ出よう(佐藤郁也 2006)』などの入門書をご覧ください。

⁵ 実際には「非標本誤差」と呼ばれる、回答拒否や誤回答など、調査をおこなう側で完全にはコントロールできないエラーが生じます。そのため全数調査を実施しても、“本来の姿”を反映した結果が得られるとは限りません。

⁶ なお「標本」とは個々のデータではなく、母集団から取り出された集団を指す言葉です。

⁷ 正確には「日本人の国民性調査」は、「層化多段抽出法」という手法で標本が抽出されています。

⁸ 標本抽出法には、非常に多くの方法があります。それらについて興味のある方は、朝倉書店刊『概説 標本調査法(土屋隆裕 2009)』などをご覧ください。

もとづいて標本を選び出す「**有意抽出法**」じゃ。確率的に標本が選出されるわけではないから、「**非確率標本抽出法**」とも呼ばれておる。また、この方法で選出された標本のことを「**有意（抽出）標本**」というんじゃ。有意抽出法は、例えば“非モテな統計数理研究所一年目の調査科学研究センター特任研究員⁹”のように抽出の基準が既に定まっていたり、または経験豊富な人がこの方法を用いる場合であれば、効率良く標本から母集団を代表する結果が得られるという利点がある。しかしその反面、思い込みや勘だけに頼ってしまうと、再現性が乏しくなるという欠点を抱えておるんじゃ。



ス：うっ、有意抽出法については分かったけれど、博士のあげた研究員の例を聞くとなぜか無性に悲しくなってくるね…。気分を変えるために、2つ目の手法について教えてよ。



ト：2つ目の方法は、母集団の中から所定の確率にもとづき、ランダムに標本を選び出す「**無作為抽出法**¹⁰」じゃ。確率的に標本を抽出することから「**確率標本抽出法**」とも呼ばれておる。そして、この手法で抽出された標本のことには、「**無作為（抽出）標本**」と呼ぶんじゃ。“確率”と聞くと、なんだか難しそうな印象を受けてしまうかもしれないが、要は「母集団に含まれている人が、すべて同じ確率のもとで偏りなく標本として抽出されるということ」を意味しているんじゃ。日本の人口が1億3千万人だとして、そこから、1万3千人を標本として抽出するとしよう。この時、全員が0.01% (=1万3千÷1億3千万)の確率で標本として選ばれる手続きがとられているならば、それは無作為抽出をおこなったということになるんじゃよ。



ス：人によって標本に選ばれる確率が1%だったり、10%だったりと異なることはないということなんだね。うん、無作為抽出法では、母集団の中にいる全員が同じ確率で標本として選出されることはわかったよ。でも、そうやって標本を抽出しても、極端な意見を持った人たちが、偶然に選出されてしまうことはあるんじゃないのかなあ？

⁹ 現在、この条件に該当するのは（おそらく）1人だけです（泣）。したがって、この基準に適合する標本を抽出すると、標本調査ではなく全数調査になってしまいます。ですので、ここでの話はあくまで例としてお考えください。

¹⁰ 無作為抽出法にもいくつかの種類がありますが、BB弾サンプリング実験のような方法を「**復元単純無作為抽出法**」と呼びます。今回の解説では、この復元単純無作為抽出について説明しています。



ト：スタッツ君がそう思うのも、無理はない。しかし、十分な大きさの標本を繰り返して抽出すると、そういった極端な特徴のある標本が選ばれる可能性は、無視しても良いくらい非常に低くなり、一方で、母集団と似た特徴を有する標本は得られやすくなることがわかっておる（＝「**中心極限定理**¹¹」）。いいかえると、適切に無作為抽出をしても、母集団と完全に一致した性質の標本ばかりが得られるわけではない。しかし、母集団ととても近い性質を持った標本をより多く得ることができるということになるのう。



ス：ああ、だから「日本人の国民性調査」では、無作為抽出法によって選ばれた人たちの意見をまとめることで、日本人の意見の大体の傾向を把握していたということなんだね。えーっと、なんとなく理屈としてはわかるんだけど、本当に母集団に近い性質の標本が得られるのか、いまいち信じられないんだよね。



ト：よし、百聞は一見にしかずじゃ。早速、BB弾を使ったサンプリング実験をして、どういった標本が得られるのか確かめてみよう！



¹¹ 中心極限定理については、昨年の解説にわかりやすく書かれています。ぜひ、そちらもご覧ください。 http://www.ism.ac.jp/events/kodomo2013/2013_sampling.pdf

2. BB 弾サンプリング実験の結果

スタッツ君の疑問に答えるため、トースター博士は BB 弾サンプリング実験をおこなうことにしたようです。実は皆さんが参加した「子ども見学デー2014」の BB 弾サンプリング実験は、その一環だったのです！果たして結果は、どうなったのでしょうか。トースター&スタッツの 2 人と一緒に見ていきましょう。

2-1. 母集団と標本の関係は？ ～標本の割合とその分布～



スタッツ：うわ～、すごい数の BB 弾だね！ 一体どれくらいの BB 弾が箱の中に入っているんだろう。



トースター：この箱の中には、母集団となる 10 万個の BB 弾が入っておる。良く見るとわかるじゃろうが、BB 弾には白玉と黒玉の 2 種類あり、白玉が 7 万 5 千個、黒玉が 2 万 5 千個あるんじゃ¹²。これを人口 10 万人の国があり、ある政策に賛成の人 7 万 5 千人を白玉、反対の人 2 万 5 千人を黒玉で表していると置き換えてやると、状況が想像しやすくなるじゃろう。このときの母集団における黒玉の割合 (=「母割合¹³ (母比率)」) は、 $25000 \div 100000$

¹² BB 弾カウンターで全ての BB 弾を数えたところ、白玉が 70023 個、黒玉は 23617 個あったそうです。このように実際の実験で用いた BB 弾と、解説で想定している BB 弾の数には差があります。これについて興味のある方は、実際の BB 弾の数で母割合を求めて標本割合を計算し、実験結果との関係を考えてみると面白いかもしれません。

¹³ 現在「割合」と「比率」では、どちらかというところ「比率」の方が多く使用される傾向があるようです。ちなみに筆者のボスは“「割合」推進委員長 (自称)”なので、本解説でも「割合」推しです。みなさんも TSK48 の総選挙 (ウソです) では、「割合」に一票を!

=0.25 じゃから、全ての BB 弾のうち 25%が黒玉ということになる。ここで無作為抽出をして、母集団と全く同じ特徴を持った標本が抽出されれば、標本中の黒玉の割合 (=「**標本割合 (標本比率)**」) もまた 0.25 になるはずじゃ。それでは、今回の実験では 300 個の BB 弾を無作為抽出することにしようかのう。ところでスタッツ君、ここで 300 個の BB 弾を無作為抽出したとき、標本に母割合と同じ割合で黒玉が含まれているならば、黒玉は何個になるじゃろうか?



ス: ええと、母割合が 0.25 だから、これと同じということは標本割合も 0.25 だよな。つまり 300 個のうちの 25%が黒玉になるわけだから、 $300 \times 0.25 = 75$ で、答えは 75 個!



ト: そのとおり。ただし先ほどいったように、無作為抽出でぴったり 75 個の黒玉を含む (標本割合 0.25) の標本が得られるとは限らん。じゃが中心極限定理にしたがえば、抽出を繰り返した後のそれぞれの標本の状態 (=「**分布**」) を見ると、75 個に近い黒玉を持つ標本が多くなっているはずじゃ。



ス: よ〜し、それじゃあ、本当にそんな標本が多く抽出されるのか、「子ども見学デー」で統計数理研究所に遊びに来てくれた人たちに手伝ってもらって、確かめてみよう!

~実験終了後~



ス: 博士、みんなに BB 弾の無作為抽出をやってもらって、95 個の標本^{14,15}を集めることができたよ!

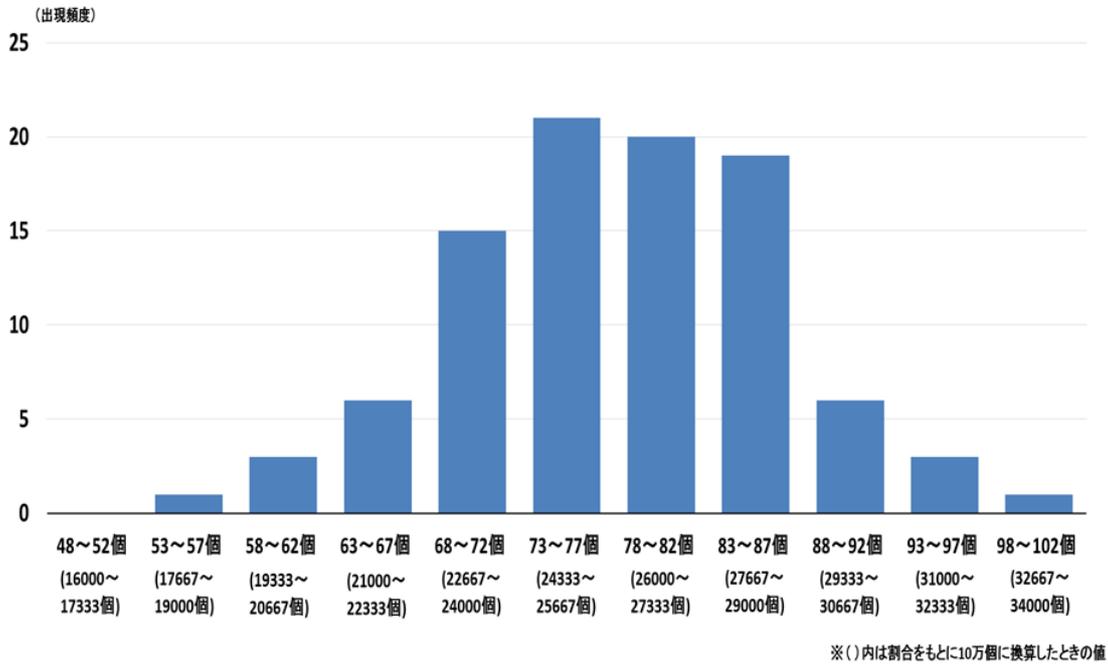


ト: よしよし、それでは何個黒玉を含んだ標本が抽出されたのか、図 (=「**ヒストグラム**」) を作って見てみようかのう。

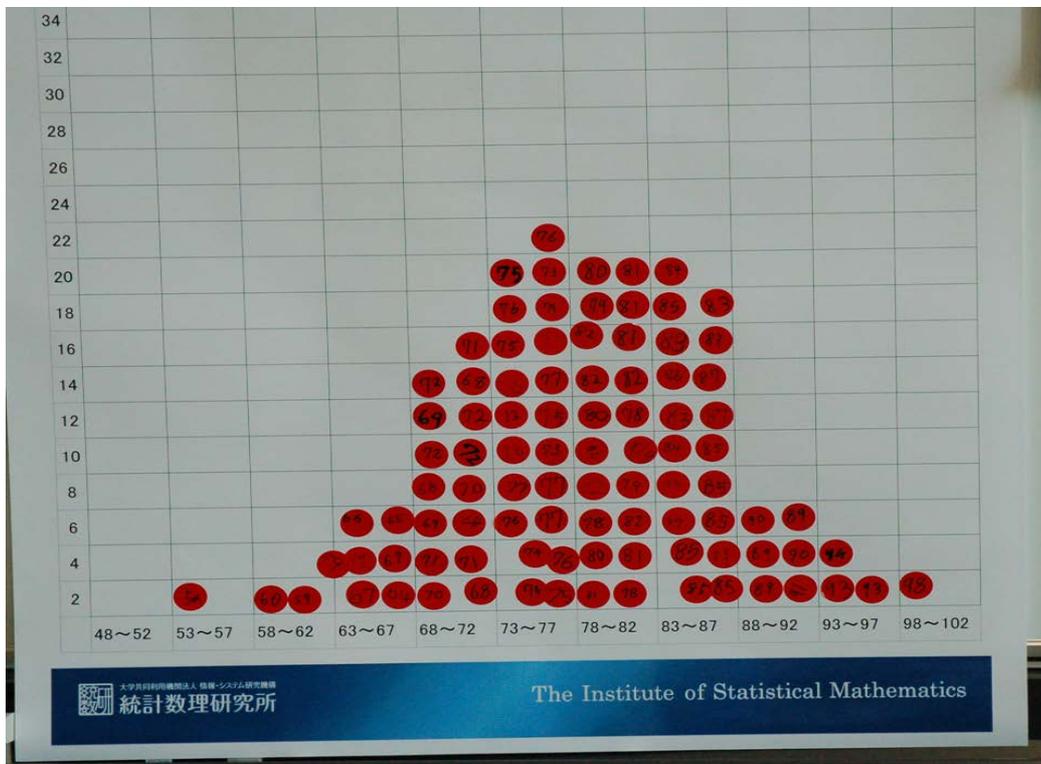
¹⁴ 標本“それ自体”の数を「**標本の数**」といい、1つの標本の中に含まれる“データの個数”のことを「**サンプルサイズ (標本の大きさ)**」といいます。今回の BB 弾サンプリング実験の例では、標本の数が 95 でサンプルサイズが 300 ということになります。なんだか、ややこしいですね。

¹⁵ 今回の 95 個の標本における黒玉の数は、次のようになりました。

74,74,81,70,98,75,68,85,60,76,85,85,76,71,71,78,89,77,91,69,80,77,68,83,59,
81,87,67,77,85,76,82,78,84,89,93,82,68,85,85,73,70,84,79,85,73,64,83,75,76
87,67,78,86,87,59,79,77,72,75,80,72,90,73,90,93,78,69,76,72,56,83,77,67,94,
83,82,82,75,65,82,85,81,72,68,79,81,80,89,73,83,65,76,84,71



実際に抽出された黒玉個数のヒストグラム



BB 弾サンプリング実験の結果写真



ス：あっ、黒玉が“73~77”個という範囲にある標本の数が一番多くなっているよ。ここでの標本割合は“0.24~0.26”になるから、この値をもとにBB弾が10万個のときの黒玉の数を計算すると、だいたい24333~25667個になるね。これって、母集団における黒玉の数2万5000個に一番近い値の範囲だよ。ということは、やっぱり母集団と似た特徴の標本が抽出されやすいんだ!



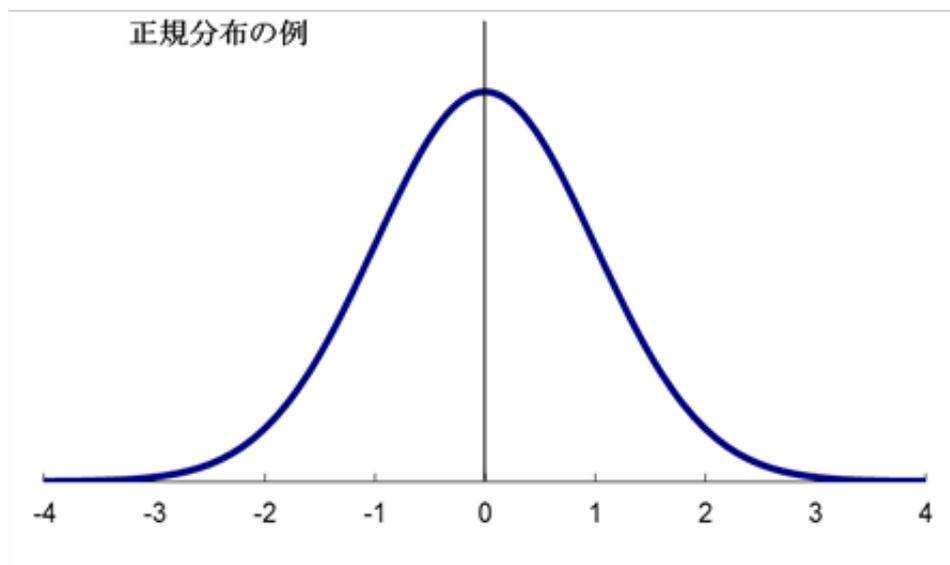
ト：それだけではないぞ。“73~77”を中心に、左右に数値がずれていくほど、抽出された標本の数は少なくなっているのがわかるじゃろう。“98~102”個の範囲で黒玉を含んだ標本は1つしかないし、“48~52”に至っては、標本は1つも抽出されておらん。これは、母集団とかけ離れた特徴を持つ標本は、抽出されにくいということを表しておる。



ス：もう1つ気づいたんだけど、ヒストグラムを良く見ると中心から離れるほど標本の出現頻度が減って行って、釣り鐘みたいな形になっているね。

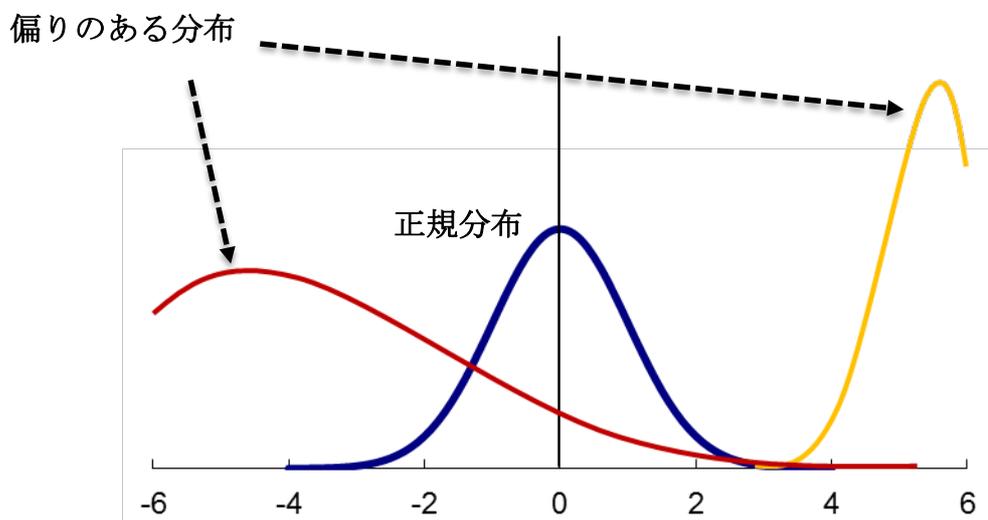


ト：そうじゃな。こんな感じで左右対称な釣り鐘型にデータが集まった状態を、「正規分布」というんじゃ。様々な自然現象のデータを集めてヒストグラムを作ってみると、実はこのような形になることがとても多いんじゃよ。





ト：さらに、得られたデータが下の図のように正規分布とかけ離れていないか調べることで、データの偏りを判断することもできるんじゃ。



ス：今回の 95 個の標本割合は、僕には正規分布のように見えるんだけど、本当に正規分布にしたがっているのかな？ それと、偏りを調べるにはどうすれば良いの？



ト：まず、データが正規分布にしたがっているかについては、「正規性検定¹⁶」という統計学の手法で調べることができるぞ。その方法を使って今回の実験データを分析すると、「正規分布にしたがっていないとはいえない¹⁷（≒ したがつている）」という、統計学的な判断を下すことができるんじゃ。ここから 95 個の標本割合は、おおよそ正規分布しているといえるじゃろう。



ト：じゃが、今回のヒストグラムを注意深く、よ〜く見てみると、母割合の 0.25 よりも、やや黒玉の割合が多い標本が選ばれてるように見えんかな？ 実はこれが、偏りなんじゃよ。



ス：そういわれてみると、そんな気がするような、しないような...ダメだ、細かすぎて僕にはわからないよ。

¹⁶ ここでは Shapiro-Wilk (シャピロ・ウィルク) 検定という手法を用いました。

¹⁷ 統計学では、まず自分が主張したいことと逆の仮説(帰無仮説)を設定します。そして、データを調べて、その仮説ではデータの状態を説明するのに統計学的な理論と矛盾が生じるとわかったときに、はじめて本来主張したいこと(対立仮説)が妥当であると考えられるのです。こうした歯切れの悪い表現になるのは、以上のような裏事情があるためです。

 ト：そんなスタッツ君のために、特別に虫眼鏡を用意したぞ。

 ス：虫眼鏡？ どこにそんなものがあるの？

 ト：その虫眼鏡というのは、「**平均値の検定**¹⁸」という統計学の手法じゃよ。今回の95個の標本における黒玉の平均値を計算すると、およそ77.8個となる。偏りを調べるには、この手法を使って標本の黒玉の平均値（77.8個）と母割合から予想される黒玉の数（75個）を比較すれば良いんじゃない。そして実際に分析すると、平均値に差がないといえない（≒ 差がある）」という結果が得られる。こうした結果をまとめると、標本割合の分布は正規分布に近いが、そこには予測よりも多く黒玉を含んだ標本が抽出される偏りが生じていた¹⁹ということがいえるんじゃない。

 ス：へえ、きちんと統計学的にデータを見ると、そんなことまでわかるんだね。

 ト：ちなみに、この偏りはBB弾のかき混ぜ方が不十分で、黒玉が箱の表面近くに集中していたことで生じたのではないかというのが、ワシの推理じゃ。

 ス：さすが博士、いつも探偵のコスプレしているだけのことはあるね！

2-2. 誤差を減らすには？ ～標本誤差、標準誤差とサンプルサイズ～

 ス：適切な無作為抽出がおこなわれれば、母集団に近い特徴を持った標本をより多く得ることができるっていうことは良くわかったよ。でも「近い」っていうことは、やっぱり標本と母集団との間に少しは差があるっていうことだよ。この差を少なくするにはどうすれば良いの？

 ト：また良いところに気がついたのう、スタッツ君。母集団と標本の特徴の差を「**標本誤差**」というんじゃない。じゃが、そもそも母集団というのとはどのような特性を持っているのかわからん場合が多いから、標本との間で標本誤差を算出することが難しい。そこで、標本の間でのバラつきを示す値（＝「**標準誤差**」）を使い、標本誤差がどの位の範囲で生じるのかを推測することになる²⁰。そして、その標準誤差は次の式で求められるんじゃない。

¹⁸ ここではt検定という手法を用いました。

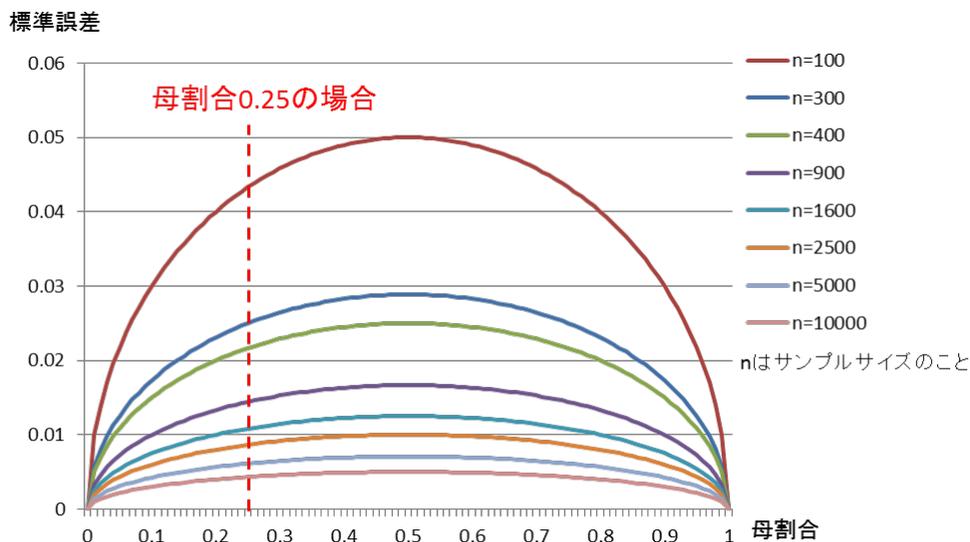
¹⁹ 正確なBB弾の数（白玉70023個、黒玉23617個）から予測される値を用いて分析しても、黒玉を含んだ標本が多く抽出される偏りがあるという結果は同じでした。

²⁰ このことについては、2年前の実験解説もあわせてご覧ください。

$$\text{標準誤差}^{21} = \sqrt{\frac{\text{母割合} \times (1 - \text{母割合})}{\text{サンプルサイズ}}}$$

ス：???、僕は数学が苦手だから、この式が何を意味しているのか全然分からないよ。

ト：数式だけでは少し難しいかもしれんな。よし、それではサンプルサイズと標準誤差の関係を図にして見てみることにしよう。



サンプルサイズと標準誤差の関係

ト：この図はサンプルサイズを増やすほど、標準誤差が小さくなっていることを表しておる。今回の実験では母割合が0.25でサンプルサイズが300だったから、赤い点線と青い曲線の交わった部分がこの実験におけるサンプルサイズと標準誤差の関係を示しておることになるな。

ト：そして、ここでもう一度標準誤差の式を振り返ってみると、平方根をとっていることに気がつくはずじゃ。このことから、精度を2倍にしたいのであ

http://www.ism.ac.jp/events/kodomo2012/20120804_sampling.pdf

²¹ ここに掲載したのは、母集団の大きさが非常に大きいことを想定して簡略化された（割合の）標準誤差の式です。また、この時の母割合は未知なので、計算をする際ほとんどの場合は標本割合で代用することになります。

ればサンプルサイズを4倍にしてやれば良いことが、なんとなくわかるんじゃないか。つまり、母集団と標本の間の差を少なくするには、抽出時のサンプルサイズを大きくすれば良いというわけじゃな (=「**大数の法則**」)。



ス：ああ、これなら僕にも理解できるよ。それにしても、小さな標本からでも、大きな集団の特徴をこんなにたくさん予測することができるんだね。だから、一部の人を対象にした「日本人の国民性調査」から、日本の中で「勤勉」、「礼儀正しい」、「親切」だと思える人が増えたっていうことがわかったんだ。ようやく納得がいったよ。



ト：そうじゃな。こうした標本調査の理論をもとに、「日本人の国民性調査」の他にも、テレビの視聴率調査や、選挙の投票率予測など、様々な調査がおこなわれておる。そのおかげで、われらは膨大な時間と費用をかけずとも、世の中の動きを大体正確に知ることができるんじゃ。そして、こうした手法の基礎となっているのが統計学なんじゃ。



ス：統計学って面白いんだね。僕、もっとたくさんを知りたくなってきちゃった。良い機会をくれて、どうもありがとうトースター博士!



ト：どういたしまして。わしもスタッツ君が統計学に興味を持ってくれて、うれしいよ。



おわりに

皆さんいかがだったでしょうか。子ども見学デーの BB 弾サンプリング実験が、どのような目的でおこなわれていたのか、そして、その結果が何を意味していたのか、おわかりいただけでしょうか。これをきっかけに、少しでも統計学や社会調査を身近に感じていただければ幸いです。

なお、「科学の道具箱（独立行政法人科学技術振興機構理科ねっとわーく <http://rikanet2.jst.go.jp/contents/cp0530/contents/>）」では、「トースター&スタツと学ぶ統計」という web ページが設けられています。統計学に興味をお持ちになられた方は、こちらもぜひご覧ください。それでは、次回の子ども見学デーで、またお会いしましょう!!

統計数理研究所 調査科学研究センター
特任研究員 稲垣佑典