

# 身長と体重における統計的分布の特性

Characteristics of Statistical Distributions of Height and Weight

物理学専攻 佐藤 勇太

Physics specialty Yuta Sato

## 0.1 研究の背景・目的

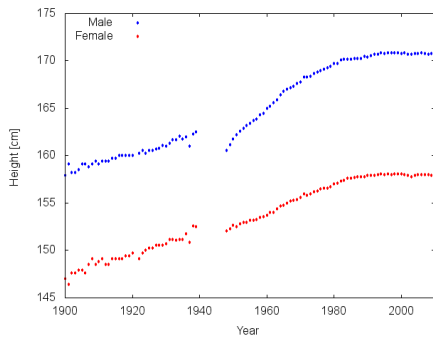


図 1: 17 歳における身長推移

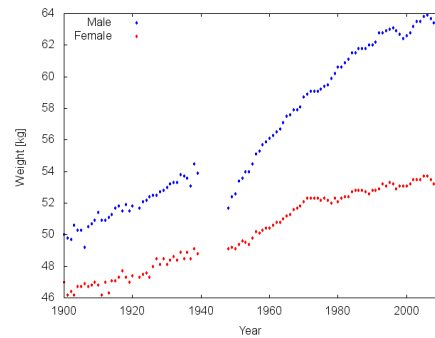


図 2: 17 歳における体重推移

図 1 と図 2 は 1900～2010 年における 17 歳男子・女子の平均身長・平均体重の推移である。ここから経済状況や生活スタイルの変化などにより、体格も常に変化しているようにみえる。実際に、戦後における男子の身長・体重に大きな落ち込みがみられるが社会や経済の立て直しとともに上昇する様子が見えてくる。

ここで示される「身長」や「体重」といった指標は健康状態をみる上でもっとも基本的なデータである。そして一般的に「身長の分布は正規分布に従う」とされている。また身長や体重は外部環境や思春期の到来などにより、正規分布に歪みが生じることが示されている。本研究では、現在の日本人の体格の様子を観察し、時間をさかのぼりながら日本人の成長を追う。そして身長や体重などの分布が時間とともにどのように変化していくかを議論し、その成長プロセスを推測する事が目的となる。

## 0.2 統計データ・データ参照方法

本研究で扱う統計データは文部科学省が公表している「学校保健統計調査」であり、幼稚園・小学校・中学校・高等学校及び中等教育学校のうち、文部科学大臣があらかじめ指定する学校に在籍する満 5 歳から 17 歳（4 月 1 日現在）までの幼児・児童及び生徒を調査対象としている。これは明治 33 年に「生徒児童身体検査統計」として始まり、戦争により中断の後昭和 23 年に「学校衛生統計」として再開し昭和 35 年に「学校保健統計調査」と名称を変え現在も調査されているものである。調査実施校については以前は単純比例抽出という方法だったが昭和 52 年度以降は各都道府県ごとに同数を抽出する確率比例抽出へ変更し、児童・生徒に関しても系統抽出という方法を用いて選定を行っている。現在は 4 月～6 月に調査実施校で行われる健康診断の結果をもとに身長・体重・座高などの発育状態と、病気の有無や視力・聴力という健康状態について調査されている。またデータ解析に伴い、男子・女子を分けて扱っている。これは男子と女子の間で成長プロセスや発育速度に違いが見られるのかを詳しく解析し、また性差による特徴的な傾向の有無の発見を目的としている。

本研究では各調査年度において年齢別に分布を調べる方法と、ある年に生まれた集団の示す分布を時間を追って追跡する方法を用いて児童・生徒の成長の様子を解析していく。本研究では前者を「横断記録」、後者を「縦断記録」と定義する。

### 0.3 解析結果 身長

図3で平成22年における17歳男子の身長の分布を示す。この図から身長の分布は左右対称で平均値を中心として釣鐘型に見え、ここから正規分布を示すということがわかる。ここでは正規分布と対数正規分布の2種類の関数でフィッティングをかけている。このままのグラフでは分布の裾の様子が分かりにくいので、 $x$  軸  $y$  軸共に対数スケールでプロットしなおしたものが図4である。裾の部分でのフィッティングがより分かりやすくなりプロットと正規分布・対数正規分布のフィットのずれ具合が鮮明になる。平均値近傍ではほとんど一致しているが、裾へ行くにしたがいずれが大きくなっていく。この図から正規分布の方がフィッティングが良好な年齢と対数正規分布の方がフィッティングが良好な年齢があることが分かり、フィッティング関数と統計データ値との残差の2乗和を正規分布と対数正規分布共に算出しどちらがフィッティングに適しているのか、また年齢によってそのフィッティング良好な関数が何かの法則に従って変化するのかを考えていく。

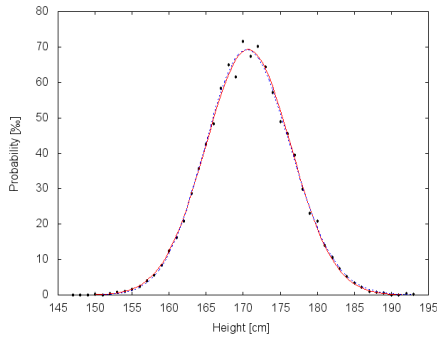


図3: 17歳男子の身長分布 (平成22年)

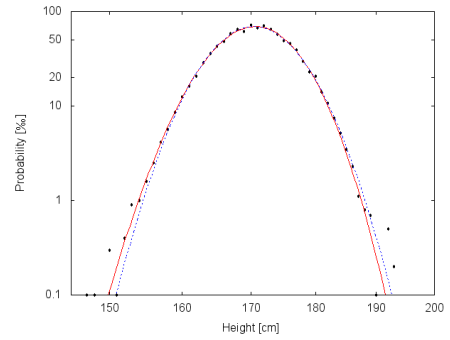


図4: 17歳男子の身長両対数分布 (平成22年)

赤の実線:正規分布、青の破線:対数正規分布でのフィッティング結果。横軸は身長 [cm]、縦軸は相対度数 [%]。

身長データを  $n$  個の階級に分けてヒストグラムを描いたとき、 $i$  番目の階級の代表値を  $x_i$ 、その度数を  $y_i$  とし、このヒストグラムに関数  $f(x)$  をフィッティングさせる時、 $i$  番目の階級の度数とフィッティング関数の値との残差の二乗和を

$$\delta_{RE} = \sum_i^n (f(x_i) - y_i)^2 \quad (1)$$

と定義し、フィッティングの優劣を確かめるために、

$$\epsilon_{RE} \equiv \log_{10} \frac{\delta_{RE}^{(LN)}}{\delta_{RE}^{(N)}} \quad (2)$$

という量を独自に定義する。この  $\epsilon_{RE}$  の大小でフィッティング関数の優劣を判定していく。 $\epsilon_{RE} > 0$ :正規分布が良いデータフィッティング、 $\epsilon_{RE} < 0$ :対数正規分布が良いデータフィッティングである。

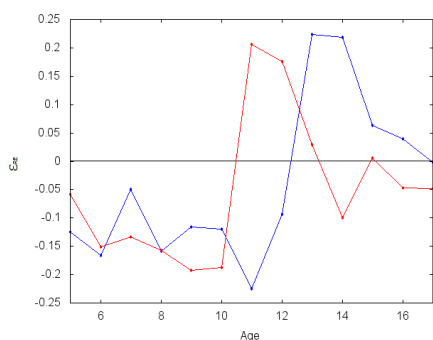


図 5: 平成 10 年 5 歳～平成 22 年 17 歳の  $\epsilon_{RE}$  推移

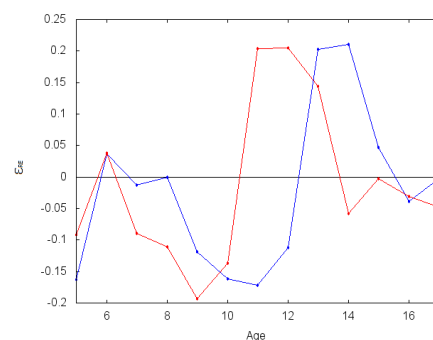


図 6: 平成 22 年における  $\epsilon_{RE}$  推移

## 0.4 解析結果 身長

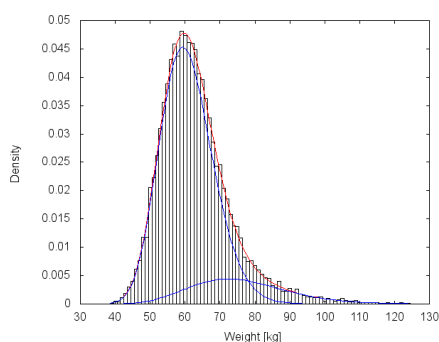


図 7: 17 歳男子の体重の分布（平成 22 年）。

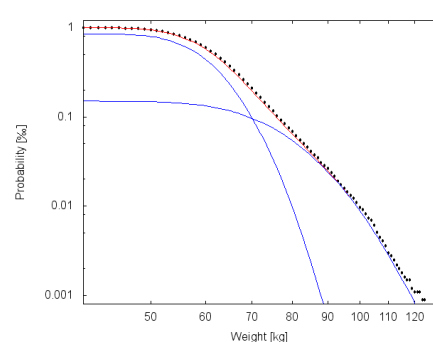


図 8: 図 7 の累積分布（平成 22 年）。

図 7 は平成 22 年の 17 歳の男子の体重分布を示している。横軸は体重 [kg]、縦軸は各階級の人数を全体の人数で割った割合とする分布である。また図 7 を累積分布形で表すと、図 8 のような概形になる。フィッティングでは対数正規分布を 2 つ足し合わせた関数を使用し、以下の式で示される形になる。

$$F(x) = \frac{N_{t_1}}{2} \left( 1 - \operatorname{erf} \left[ \frac{\ln(x/\mu_1)}{\sqrt{2}\sigma_1} \right] \right) + \frac{N_{t_2}}{2} \left( 1 - \operatorname{erf} \left[ \frac{\ln(x/\mu_2)}{\sqrt{2}\sigma_2} \right] \right) \quad (3)$$

このように対数正規分布を複数構造とすると、体重の分布は非常によりフィッティングを見せる。

昭和 45 年の段階ですでに累積分布の重ね合わせとして多層化の傾向が見られているが、年を経るごとに体重の大きい集団の厚みが増加し分布の裾が広がっていく傾向がある。しかしいずれの場合も思春期以降、体の成長が止まる年齢のあたりで多層化していた対数正規分布のうち 1 層が減少する様子が伺える。分布の形状は平成 22 年では男子・女子共に 5 歳で 3-lognormal distributions で始まり、男子では 12 歳で 2-lognormal distributions に移行しその後は 17 歳まで 2 層化のままである。女子では 10 歳で 2-lognormal distributions に移行し、13 歳で再度 3-lognormal distributions へ移行する傾向が見られる。昭和 28 年の体重分布は男子では 9～14 歳で 2-lognormal distributions それ以外は 1-lognormal distributions、女子では 11～12 歳で 2-lognormal distributions それ以外は 1-lognormal distributions でフィッティングされる。

出生した段階では体重の分布は複雑化しておらず、成長するに従い分布が2層化、3層化の構造を持つようになる。構成している複数の対数正規分布の層が減る際には、2つの対数正規分布の中央値  $\mu_i$  が接近するかデータ数  $N_{t_i}$  が相対的に小さくなる。2つの分布を1つにまとめて簡略化した方がフィッティングが良くなる傾向にある。

## 0.5 まとめ

本研究で知り得た事を簡潔に以下にまとめる。

### ・身長

平成22年の男子・女子の身長の分布は、思春期前は対数正規分布、そして思春期の到来とともに正規分布へ移行し思春期後は  $\epsilon_{RE}=0$  の成長が収束へと向かう様子が見られる。このことから思春期前の段階ではジブラ過程に従い成長し、成長プロセスの変化を迎え正規分布型の成長過程になる事が見てとれる。昭和28年・昭和45年・平成3年・平成8年・平成13年・平成18年の横断記録や、平成10年の5歳を基準とした縦断記録からも同様の傾向を見ることができる。身長の成長は近年の平均身長の伸びというものはあるが、プロセス自体は普遍的なものであると考えられる。

### ・体重

平成22年の男子・女子の体重の分布は、いずれも思春期の到来とともに構造が3層から2層または2層から1層とシンプルになる傾向が伺える。これは近年に見られる傾向であり、昭和28年の分布では5歳で1層の関数が思春期で2層になりその後1層に戻る様子が見られる。ここから思春期前の多層化は近年の生活環境の改善によるものであると考えられるが、思春期の到来とともに層が減少し簡略化することは普遍的な傾向であると考えられる。平成22年の女子の場合は単純ではないが、いずれにしても複数の対数正規分布の累積分布形の和で表す関数が、体重の分布においてはフィッティングに有用な場合が多いことが分かる。

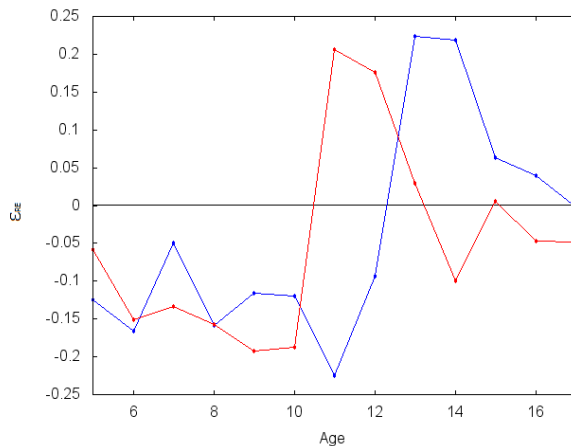


図 9: 身長における残差の2乗の傾向

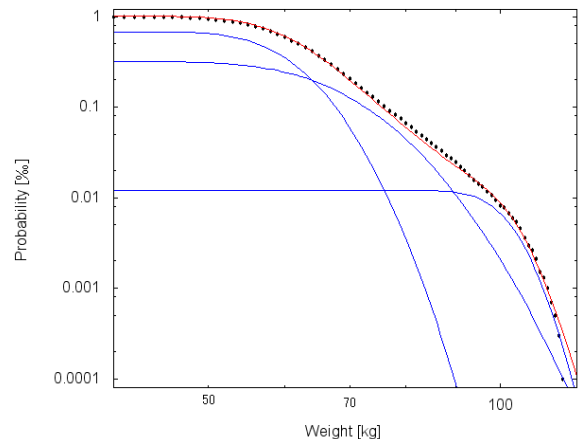


図 10: 体重における関数の重ね合わせ