



e-数字を使える営業マンは仕事ができる

—数字嫌いのための統計活用術—

◆ 初版 : 2003年9月

◆ 発行所 : 日本経済新聞社

使用上の注意

- 本テキストは、ライセンスを受けた本人のみが閲覧可とします。
- 本テキストは、ダウンロード後、いかなる媒体へのコピーもこれを不可とします。ただし、ライセンスを受けた本人が個人的な使用のため、一度だけプリントアウトすることを許可します。
- 他の媒体への引用、転載は一切認めません。

私は1年半ほど前に日経文庫から「マネジャーが知っておきたい経営の常識」という本を出版しました。マネジャーとして知っておくべき知識を体系的にまとめたものであり、おかげさまで好評で増刷を続けてきました。読者層も当初想定した方たちよりも幅広く、新入社員から経営者まで色々な方に読んでいただきました。しばらくして読者の方から、「経営、マーケティング、組織などはよくわかるが、最後の章にある『将来を予測する』のところがよくわからない。何か良い参考図書はないか」という問合せを数多くうけました。

私はコンサルタントとしてクライアントに「これからのビジネスは、売上を最大にしていくことよりも、顧客の動きをとらえ、明日の売上を読むことだ」と言い続けてきました。しかしよく考えてみると、「明日を読む」ことにおいてその考え方や具体的方法論を提供していないことに気づきました。

「将来を予測する」ということを勉強するための、適当な書籍を探してみたのですが、良いものが見つかりません。ほとんどが数学の専門書か、表計算ソフトの統計機能の解説書です。前者は数学者が書いているため、理論的にはしっかりしているのですが、例がサイコロやコインなどであり、ビジネスケースを扱ったものはほとんどありません。数学者がビジネスを体験することがないからでしょう。

この本を読んだ多くのビジネスパーソンは自らの数学的素養のなさをなげき、半分も読まないうちにギブアップしてしまいます。何とか最後までたどりついた勉強家も確率や統計を理解できても、具体的にビジネスにどう使っているのかわかりません。

後者の本ではビジネスのケースが多いのですが、理論的バックボーンがあまり解説されておらず、これを入れればこれが出ると書いてあるだけです。そのため表計算ソフトの統計機能を使うことができても、そこから出てくる数字を上司や顧客に説明することができず、結局使われなくなっていくます。

1年くらい悩みましたが、私が出した結論は自分でそのニーズに答えることです。私は数学者ではありませんが、小さい頃から数学は大好きでした。今でも受験勉強中の高校生の長女から自分で解けない数学の問題を質問されるとワクワクしてしまいます。そして長女には「数学は式を覚えることではなく、問題を自分で考え、解き、そして先人たちが考えた解き方を学ぶんだ。解き方を覚えて、そのトレーニングをすることではない」といつてきました。これと同じ言葉を私が指導している企業のセールスパーソンに言ったことが本書を書こうと思ったきっかけです。

本書の目的はこの後のプロローグに書いてあるとおり、ビジネスパーソンが自分の考えたことを「数字で説明する」力を身につけることです。数字の使い方を体系的に整理した学問が数学であり、それを学ぶことで数字に強いビジネスパーソンになってもらうことが本書のねらいです。そして対象は「数学の苦手なすべてのビジネスパーソン」です。読んでいただき、ビジネスへの数学の活用方法を理解していただくとともに、少しでも数学の楽しさを味わっていただければ幸いです。

2003年8月 MCシステム研究所 内山 力

がんばります

「部長、来期の売上見通しを作りました」

「どういう基準で作ったんだ」

「各顧客すべて今期の5%アップにしてみました」

「5%の根拠は何だ」

「・・・特にはないです。私としてはできれば2%くらいにしたいところですが、5%くらい上げないと会社としてはまずいでしょ」

「何言ってるんだ。そんなことは経営側で考えることだ。実際に商売をやっている君たち現場から見て、どの位の数字が妥当かを知りたいから聞いてるんじゃないか」

「うーん、こんな時ですので来期はちょっと厳しいですね。ライバルの手が読めないんで、なんともいえないんですが、まあ普通にやれば今期より少し落ちるでしょうね」

「少しってどれ位だ」

「やってみなければ何ともいえません。まあ、やれといわれれば強気の計画でも何でも組みますよ。もちろん体を張ってがんばります」



日本中の企業で見られるシーンである。希望的数字なのか、予測した数字なのか混乱し、数字だけが一人歩きしていく。そして数字だけがんばってしまう。プロセスをがんばるはずが、目標設定でがんばってしまう。

せっかくのデータがドブに

「お宅の商品売れないから、うちの売場からカットするよ」

「えー勘弁してくださいよ。他の店では売れてますよ。陳列の仕方だと思いますよ。ほら、これがそのレポートです」

「うちと客層がちがうんじゃないの。どこの店のレポートなの」

「ニュースソースは勘弁してください」

「売れない店のデータをカットして、売れている店のデータだけを持ってきてんでしょ。これだからセールスは怖いよな」

「そんなことはありません。お宅でもきっと売れますから」



自らの意見と事実としてのデータが混乱しており、データの分析も自分にとって都合のいい結果だけを集めて顧客に提出している。そのため逆に顧客から信頼を得られない。企業が一生懸命お金をかけて作ったデータを、セールスパーソンが皆自分の勝手な色をつけ、別のデータに変えて、そして顧客から信頼を失い、結局それをドブに捨てている。

コンピュータを妄信

(部長が、レポートを作成した情報システム部に)

「何だこの来期の売上予測は？いくらなんでも数字が大きすぎるだろう。レポートを作った情報システム部にどうやって出したか聞いてみよう」

「これは重回帰分析で求めています。説明変数としては人口、世帯数…」

「うーんまいった。よくわからないけどコンピュータがやったんなら信じよう」

(現場担当の部下が部長に)

「部長、この来期の売上予測値、どうやって作ったか教えてください。こんなに大きくなるとは思えませんが…」

「そんなことはどうでもいいだろう。数字をどうやって作ったかより、現場としてはいかにそれを達成するかだろう。精一杯やってみろ。やる前にとやかく言うな」

「でもこんな数字の達成度でボーナスまで決まるんですか？」



どんなに良い方法で予測しても、その数字がどうやって出たのか、使う人が理解してなければ何の意味もない。誰か数学が得意で、コンピュータが好きで、現場を全く知らない人が作った数字が、いつの間にか会社中を一人歩きしてしまう。

「数字よりも気力」になってしまうわけ

これだけ日本中が数字に弱ければ、数字をうまく使ったトヨタ、セブンイレブン、花王などが業界で1人勝ちとなってしまう。なぜ日本の企業では数字がうまく使えないのだろうか。筆者はどうも学校教育で理系と文系に分けることにあるように思う。この分岐はほとんどの人が数学が「好き」か「嫌い」で決めているようである。それほど数学は「好き」か「嫌い」かがはっきりしている学問といえる。数学は論理性のかたまりのような所があり、直感的、非論理的な人を一切受け入れない「かたくなさ」がある。

そして人は一度数学をきらいになると、二度とその考え方を受けつけないようになっていく。こうして分かれた理系、文系は一生つきまとう。就職時にはっきりと区分され、多くの企業では別のキャリアステップを進んでいく。理系の就職先（企業のことだけでなく職種のこと）は工場、研究所、エンジニア部門、情報システム部門であり、スタッフ志向が多い。文系は企画、セールス、広報、経理、人事…などそれ以外のすべての部門であり、ライン志向が多い。こうして組織を作っていくと理系が少ない（そもそも学生が少なく、また就職せず学問を究める人も多い）企業では、経営者たる取締役が全員文系のこともある。そしていつの間にか数字よりも気力を大切にする企業文化ができあがっていく。

「ビジネスに使える数字」とは

数学は工場や研究所には活かされている。本書で扱う主な分野である確率、統計は工場の品質管理とともに発展したといってもおかしくない。

しかし、マーケティング、経営にはほとんど活かされていない。デフレ時代を迎えた今、企業に求められているのは、爆発的なヒットではなく、売れるもの、売れる量、売れる時期を予測する力であることは異論のない所だと思う。そしてその需要予測を担うべき人材はセールスパーソンなど第1線で働くビジネスパーソンであり、決してビジネス現場を見たこともない統計の専門家ではない。

本書はこのビジネスパーソンに先人の知恵ともいえる数学、とくに予測のための統計学を身につけてもらうことを目的としている。ビジネスに数学を活用するうえで大切なことは「どうやって予測するか」といった「やり方」そのものではなく、「なぜそういうやり方を使ったのか」という説明力である。やり方はコンピュータに覚えさせれば良いし、もっと言えばブラックボックスでも良い。

自らの目で数字を見て、そこから得られる自分のカンを大切にして、そのカンを活かして、数字を加工し、表現していく力である。その数字を加工・表現して

いくために先人のやり方を学ぶという姿勢が大切である。

自らのカンを活かし、自らがやり方を「選べば」（「作る」のではなく）、他人にその数字が出たプロセス、意味をきちんと説明できる。そしてその時、その数字は生き物となり、上司やお客様と話し合っていく中できつとビジネスに使える数字へと成長していくはずである。

ビジネスに役立つことがわかっているのに、なぜ文系のビジネスパーソンは統計学を勉強しないのだろうか。特に将来の数字を予測することが仕事ともいえるセールスパーソンは近づくことさえいやがる。彼らと実際に話してみてもやっとわかった。それはまえがきにも書いたが、良い教科書がないのである。本屋に行けば統計の本は山ほどあるが、この本は2つに分類される。1つは数学者の書いたバリバリの純粋数学本である。微分、積分なんて忘れ、関数さえも忘れた（中学でやったことも忘れ）ビジネスパーソンは式を見た瞬間に本を閉じてしまう。もう1つはいわゆるテクニク本である。表計算ソフトなどを使ったやり方が書いてあり、「これを入れればこれが出る」というものである。確かに数字は出るがこれでは他人に数字を説明できない。

「考え方」を理解する

本書は数学を「やり方」でなく「考え方」を中心に解説してある。主な学習項目は次のとおりである。

第1章…母集団、標本、平均、標準偏差

第2章…需要予測モデル、指数平滑法、時系列分析、最小2乗法

第3章…回帰分析、相関係数、多変量解析、数量化理論

第4章…EOQ、確率分布、正規分布、安全係数

第5章…区間推定、t検定、F検定、カイ2乗検定

具体的に数学を必要とするビジネスシーンを提示し、そこで使わざるを得なくなったビジネスパーソンが四苦八苦しなから本を読み、人に教わり、学習していく姿を書いている。そしてその姿を追いかけていくことで、自然にこれだけの項目が身につくようになっていく。

また本文のストーリー部分では数学の知識がなくても「やり方」の本質が理解できるよう工夫している。この「やり方」に加えて微分、積分といった本文中に出てくる数学自体をしっかりと理解したい人のために、章末にコラムを入れて、ここで詳しく解説している。

本書は数学が苦手なセールスパーソンである佐藤氏が次第に成長していく姿を書いたものである。読者の方々は彼と自らの姿をオーバーラップさせ、「他人に数字が説明できる」ビジネスパーソンへと成長して欲しい。

目次

第1章 商品力を数字で表す

お客様に必要とされる営業マンになりたい

うちの商品には「力」がない？

グラフなんて役に立たない？

本を読んでもチンプンカンプン

視聴率と売上予測に共通すること

「明日」は過去の延長線上にある

まず知りたいことを決める

調べる対象をはっきりさせる

過去と将来の「同じ部分」と「違う部分」

「わかっている数字」と「わかっていない数字」

「傾向」に目をつける

原因を考えるのは人間の仕事

商品力って何？

増加率の平均の出し方

元の数字に戻れ

リピート率が高い=コンスタント

バラツキを調べる

2倍のコンスタントさ

数字から導き出されたセールスポイントは？

コラム:期待値

第2章 明日はいくつ売れるかを予測

顧客と目標を共有する

普通にやれば売れる数字

時間が関係するデータ

最適な方法の見つけ方

トレンドが見たいけど・・・

「たまたま」を取り除く

前期の予測値と実績値で予測する

過去の実績をすべて使う

カンを式に生かす

もっと売れるはず

なぜそういう予測をしたかを残しておく

弱気すぎる数字

商品のライフサイクル

成長期には使えない手法

点を打っていく

線を目分量で引く

「人に説明できる形で線を引く」とは？

どうやってコンピュータに線を引かせるか？

予測値が納得できないとき

先の読める人間

これが「普通にやれば売れる数字」

コラム：微分

第3章 新しい店の売上を考える

オープン前なのに、年間の売上高を予測？

他店のデータから予測する

「戻る」分析って何？

「こづかい」と「年収」

標本数が違えば式が変わる

「線」が見えるか見えないか？

「これが回帰分析だ」

説明変数がたくさんあって決められない！

相関分析は回帰分析とどこが違うの？

直観を数字で表す

どのデータで予測するか

説明しやすいものを選ぶ

線を引いて式を探す

限られたデータなら直線でOK

「箱」の中で線を引く

目的は数字の意味を説明できるようにすること

売上をどれくらいと考えるのが妥当か

ありが1でなしが0

予算や顧客分析にも使える

コラム：指数と対数

第4章 発注ロットを考える

発注の小口化で物流コストが増加

コンビニが発注ロットを小さくするわけ

総費用を最小にするロット

買う方は小さく

売る方は大きく

売る方と買う方の利害が一致するポイント

を考える

欠品と売れ残りをなくしたい

身体検査で身長を測ることと確率の関係

ヒストグラムで考える

棒の頭をつないでみる

面積が確率

つり鐘型のグラフ

世の中の現象は「つり鐘型」のものが多い

一つの表ですべてわかる

欠品を出さない在庫数は？

10回に1回の品切れ

コンスタントに売れると在庫が減る

ROIはビジネスの基本

2日分の安全在庫は2倍にならない

コラム:積分

第5章 数字で説得する

クロスマーチャンドアイジングの試み

実験は成功させよう

トマトの売上が1.5倍に

売上が伸びたのはフェアで客数が増えたから？

値下げや販促も原因では？

実験やり直し

やはり売上は伸びた

たまたまじゃないの

難攻不落な人を納得させる説明とは？

まず反対のことを仮定する

「トマトソルトを置いても効果なし」と考える

仮定した結果になる確率を出す

「効果なし」は誤り

フェースを2倍にしたら売上も2倍？

「客の多くが入り口から左に行く」というのは

正しいか？

「等しい」ことが「否定できない」=「等しい」？

バラツキを見るF検定

平均値の検定の前にバラツキの検定を

仮定と現実が合っているかを検定

リーダーへの道

コラム:確率分析

エピローグ

第 1 章

商品力を数字で表す

佐藤一郎はミドリ食品という食品メーカーのS支店に勤める営業部員だ。

4年目に入った今年から、MCスーパーの担当を任されることになった。ミドリ食品では若手営業マンに通りの商品知識を身につけさせると、まずBランク*の顧客の掘り起こしをやらせるのが常だった。

MCスーパーはS県を中心に40店舗を展開する地域密着型のスーパーだが、ミドリ食品が今注目している健康飲料の分野ではライバルメーカーのネイチャーフーズに最近押され気味だった。

佐藤「よし、がんばるぞ。ただ単に商品を売るだけでなく、お客様に必要とされるセールスになろう」

Bランク

顧客、商品などを売上、粗利額などによって3ランクに分類するものをABC分析という。メーカー、流通業では一般的に顧客を取引高に応じてAランク（上位20%の優良顧客）、Bランク（20～50%の中位にある顧客）、Cランク（それ以外）に分けて管理している。

佐藤は事前調査のため、MCスーパー中央店に買い物に行った。

店内は地元の主婦でにぎわっていた。健康食品売場は、折からの健康食品ブームで主力の売場の一つに育ちつつあったが、残念ながらネイチャーフーズの商品が大きなスペースを占め、ミドリ食品の商品は探すのも大変だった。

翌日、佐藤は上司の中村支店長、前任者山田とともに、MCスーパー中央店の店長工藤を訪問した。店長は先月まで別の店の生鮮部門のバイヤー*だったという。

中村支店長は店長に佐藤を紹介し、ミドリ食品が今力を入れている「スーパーグリーン」というコレステロール値を下げる効果がある健康飲料の売上増に協力をお願いした。

バイヤー

スーパーなどでの商品の仕入担当者をさす。中規模以上の店舗では、加工食品、生鮮食品、日用雑貨などの部門単位におり、当該部門の専門的知識と購買決定権を持っている。

工藤 「スーパーグリーン？ ああ聞いたことあるけど、うちにおいてあったかなあ。まだ来たばかりでよく把握していないんだ。加工食品のバイヤーやっている木村に聞いてみるよ。ちょっと、木村君」

木村 「はい、何でしょう」

工藤 「今、ミドリ食品さんが来てるんだけど、スーパーグリーンって調子はどうだい」

木村 「はあ、前に山田さんに言われて、フェース*増やしたけど、あまり出ませんでした。やっぱりネイチャーフーズのナショナルヘルスシリーズには勝てないんじゃないでしょうか。いくらメーカーのセールスががんばっても、売れ行きを左右するのは価格と商品力でしょう。健康飲料は価格をいじっても動かないから、今の定番*で固めるしかないですよ。フェースの入れ替えはむずかしいですね」

中村 「おかしいなあ。他の店では結構売れているのですけどねえ。ここ3ヶ月の売上データを見せてもらえませんかね？」

木村 「店長さえ良ければ私がかまいませんよ。POS*データは渡し方がよくわからないけど、商品別の月報ならありますよ」

中村「佐藤、こう言っていたんだから、それをお借りして、しっかり分析してみろ。店長、木村さん、ありがとうございます。必ずMCスーパーさんの売上を伸ばすような提案をさせていただきます」

佐藤に渡されたのは、健康飲料売場に置いてある商品の、4月第1週からの3ヶ月間の毎日の売上個数、金額が記録されている表だった。(図表1-1)

6月月報

5月月報

4月月報 (15.04.07~15.05.04) 中央店: 健康飲料

商品名	7日(月)	8日(火)	9日(水) ……
スーパーグリーン	2 (500)	3 (750)	5 (1,250) ……
ナショナルヘルス (ハイパー)	6 (1,800)	10 (3,000)	2 (600) ……
ナショナルヘルス (ナチュラル)	4 (1,200)	8 (2,400)	6 (1,800) ……
ナショナルヘルス (365)	7 (1,050)	1 (150)	4 (600) ……

()内は売上げ

図1-1

 フェース

フェースとは店舗での陳列に関する用語で、来店客から見てその商品が何個見えるかということを意味する。棚で言えば1列という意味。陳列数はフェース×奥行き数（1フェース当たり何個置くか）となる。

定番

定番商品のこと。季節や店舗でのプロモーション（特売など）に左右されず常備している商品。また店舗内で常設されている売場をさすこともある。

POS

Point of Salesの略で、本来は販売時点で情報を管理するという意味。一般には、バーコードなどを読み取るリーダーを持つレジスターによって迅速にキャッシングを行い、そこで発生したレジデータ（POSデータという）を店舗運営に活用していくことを言う。

会社に戻った佐藤は、前任者の山田に相談してみた。

山田「こんなこといつもやってるよ。俺が表計算ソフトで作った表に入れば一発でOKだ。きれいなグラフも出るぞ」

佐藤「パソコンで表を作るのはあまりやったことがないのですが…」

山田「大丈夫。箱の中に数字を入れていくだけだよ。ただ商品は絞った方がいい」

佐藤「じゃあ、スーパーグリーンとナショナルヘルス3種の4つでやってみます」(図1-2)

		第1週								
		4/7(月)	4/8(火)	4/9(水)	4/10(木)	4/11(金)	4/12(土)	4/13(日)	週合計	週平均
スーパーグリーン	個数	2	3	5	4	6	8	12	40	5.7
	売上額	500	750	1250	1000	1500	2000	3000	10000	1428.6
ナショナルヘルス(ハイパー)	個数	6	10	2	4	2	18	16	58	8.3
	売上額	1800	3000	600	1200	600	5400	4800	17400	2485.7
ナショナルヘルス(ナチュラル)	個数	4	8	6	8	0	6	14	46	6.6
	売上額	1200	2400	1800	2400	0	1800	4200	13800	1971.4
ナショナルヘルス(365)	個数	7	1	4	2	10	14	18	56	8.0
	売上額	1050	150	600	300	1500	2100	2700	8400	1200.0
		第2週								
		4/14(月)	4/15(火)	4/16(水)	4/17(木)	4/18(金)	4/19(土)	4/20(日)	週合計	週平均
		3	4	6	5	7	9	13	47	6.7
		750	1000	1500	1250	1750	2250	3250	11750	1678.6
		4	5	9	1	7	16	12	54	7.7
		1200	1500	2700	300	2100	4800	3600	16200	2314.3
		14	2	13	3	6	25	4	67	9.6
		4200	600	3900	900	1800	7500	1200	20100	2871.4
		0	6	8	1	3	6	2	26	3.7
		0	900	1200	150	450	900	300	3900	557.1

図1-2

佐藤は山田から渡されたファイルに4種類の商品の数字を入れ、作ったグラフを支店長に見てもらった。

佐藤「支店長、先日のMCスーパーのデータ分析ができました。いかがでしょうか」(図1-3)

ところが、中村支店長は一目見るなり突っ返してきた。

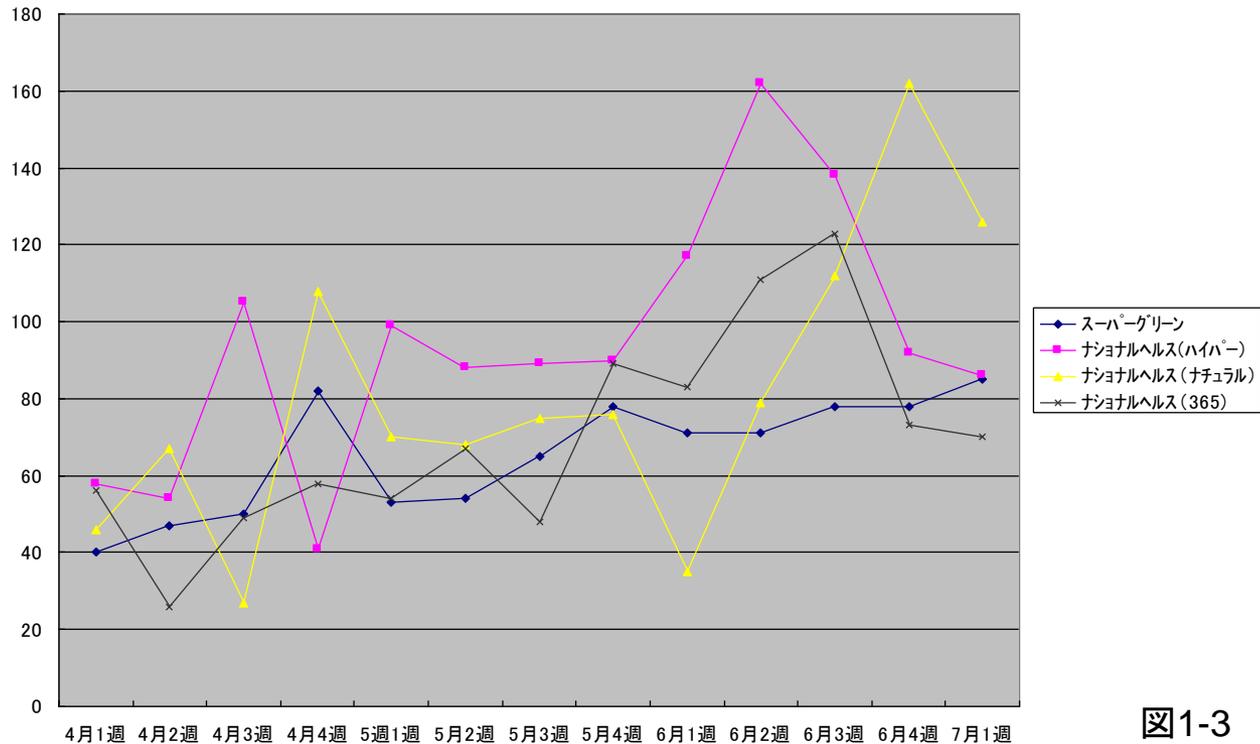


図1-3

中村 「ただ数字をグラフにしただけじゃないか。これをあちらに持って行って何というつもりだ。大体こんなグラフ、向こうも持っているぞ。POSシステムで普通出てくるだろう」

佐藤 「でも、グラフを作って気づいたんですけど、4月の4週目のゴールデンウィーク直前では、うちのスーパーグリーンが、ネイチャーフーズのリーダー商品のハイパーに勝っています」

中村 「だから？そんなこと、MCスーパーさんに言ってどうなるんだ」

佐藤 「はあ……」

中村 「頭を使えよ。どうしたらよいか自分で考えろ。あちらもあせっているわけじゃないから、一度数字の使い方をしっかり勉強してみろ」

佐藤 「数字の使い方？」

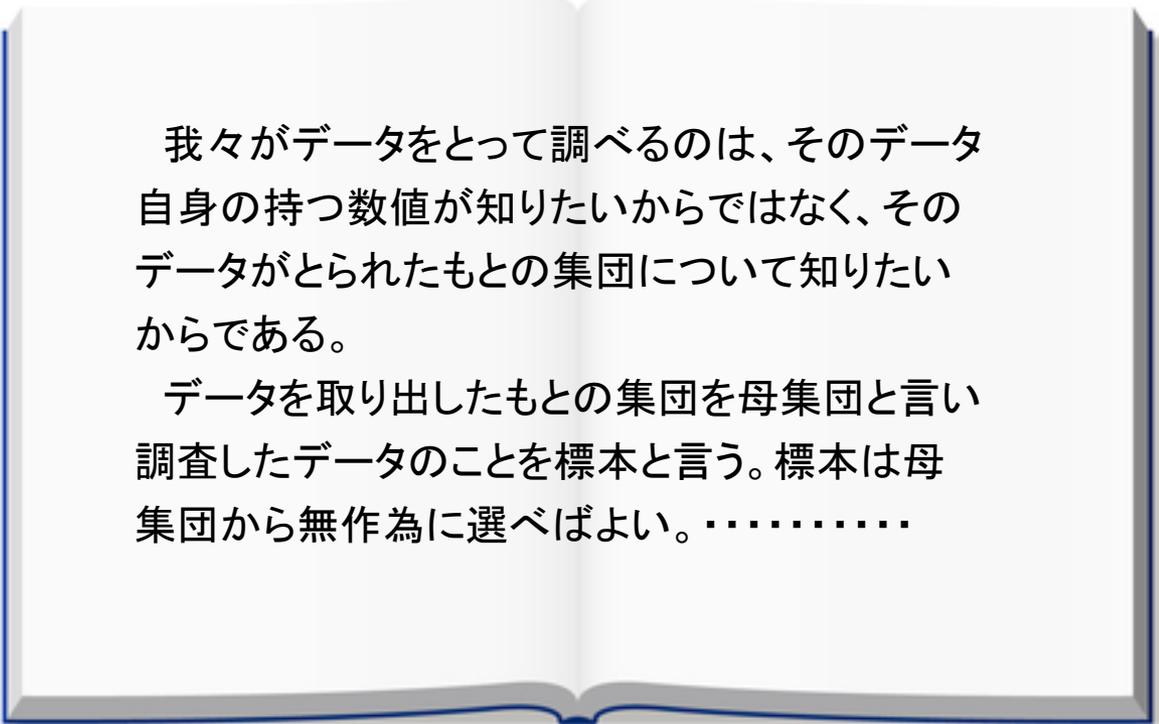
中村 「数字をうまく処理することを統計学というんだ。聞いたことあるだろう」

佐藤 「統計学って数学ですよ。数学が一番苦手なんですけど……」

中村「仕事に必要なんだから、好き嫌い、得意・不得意なんて言っていられないだろう。これだけモノが売れない時代のセールスは、売れる量を予測することが最も大切だ。それが統計だ。数学なんて思うから難しいんだ。セールスマニュアルだと思って本を読んでみる。どうしてもわからなかったら販売企画課の土屋に聞け。彼はこの道20年のプロだ」

佐藤はインターネットで検索して、「ビジネスマン向けで最もやさしい統計の本」と書いてある『誰でもわかる統計の本』を買った。そして1ページ目の「母集団」「標本」「無作為」という言葉で止まってしまった。(図1-4)

仕方がないので販売企画課の土屋のもとを訪ねた。



我々がデータをとって調べるのは、そのデータ自身の持つ数値が知りたいからではなく、そのデータがとられたもとの集団について知りたいからである。

データを取り出したもとの集団を母集団と言い調査したデータのことを標本と言う。標本は母集団から無作為に選ばばよい。……………

図1-4

佐藤 「中村支店長から土屋さんが統計に詳しいときいて、うかがいました。母集団って何ですか？標本って何ですか？」

土屋 「へえ、この支店にもそんなことを勉強しようと思うセールスがいたんだ。感動だなあ。2年前にこの支店に来て、最初にデータ分析の勉強会をやったんだ。そしたらものの30分も立たないうちに、皆寝ちゃったもんなあ。それっきりやる気なくしてたんだよ。君みたいな若いうちに、こういうことを勉強すれば必ず身につくよ」

佐藤 「すみません。自分は中学・高校と数学はまるっきりだめなんです。大丈夫でしょうか」

土屋 「心配するな。俺も話し相手が出来てうれしいよ。俺はもともと工場出身で、社長からマーケティングにデータを活かせといわれて販売企画課に来たのに、誰も相手してくれなくて、何をやっていいのかわからず、困ってたんだ。ところで、質問は何だっけ？」

佐藤「母集団と標本です。本を読んだのですが、出ている例がテレビの視聴率で今1つピンと来ないんです」

土屋「それはな、全国の家で紅白歌合戦を何人見たか知りたい時、全家庭に電話して聞くと大変だから、あらかじめ決めておいた一部の家庭に機械をつけて、何時から何時までどんな番組を見ているかを調べて出すのが、視聴率だろう。この全国の家が母集団にあたり、機械がついている家庭が標本、つまりサンプルにあたる、と書いてなかったか」 (図1-5)

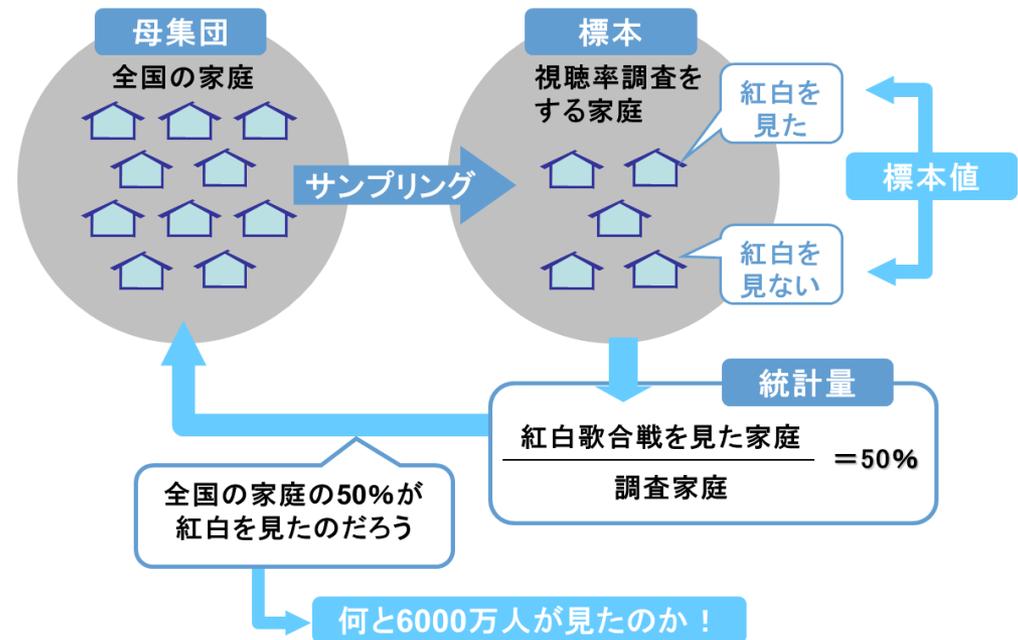


図1-5

佐藤 「標本値という言葉と統計量という言葉が出てくるのですが、違いがよくわからないんです」

土屋 「標本が家庭なら、標本値は紅白を見たとか見ないという標本を表しているデータだ。コンピュータでは見ないと0、見ると1だ。統計量とはこの標本値を加工して作った計算値のようなものだ。代表例が平均値や、学校の試験なんかで使う偏差値だ。ここでは視聴率が統計量だ。まあ母集団の特徴を表すようなものだ」

佐藤 「もう1つわからないのは、期待値という言葉です。平均値と同じだと思うのですが？」

土屋 「数学者は結構細かくて、言葉の定義を大切にするんだ。そこが文系と理系の違いだな。平均値というのは過去の結果だ。期待値というのは明日期待される値だ。商品がこの4日間で5個、6個、7個、6個と売れたら、平均値はいくつだ？」

佐藤 「6個です」

土屋 「期待値は？つまり明日何個売れると考える？」

佐藤「6個です。そうか、平均値を使って期待値を出すんだ。最後に、無作為って何ですか」

土屋「ランダム of 日本語訳だ。標本の選び方をサンプリングという。日本語でいう抽出だ。要するに標本をサンプリングするとき、思いを入れないことだ。例えばNHKの関係者は紅白の視聴率が高くなって欲しいよな。若い人が紅白を見ないなら、その若い人がチャンネル権を持っている家庭は標本に入れたくないよな。これが作為だ。そして何も考えずに標本家庭を選ぶのが無作為だ」

佐藤 「そこまでは何となくわかりました。でも今回のMCスーパーの時はどう考えていいかわからないんです。MCスーパーさんからこのようなデータをもったのですが…」

とって、4月1週から3ヶ月間の月報を見せた。

土屋 「君はこのデータで何を知りたいんだ？」

佐藤 「えーっと、MCスーパーのバイヤーが言っていた商品力というのを考えてみたいんです」

土屋 「商品力って何だ？」

佐藤 「消費者が商品を買いたいと思う度合い…でしょうか…」

土屋 「それなら視聴率と同じだろう。何%位の人が買いたいと思うかだろう」

佐藤 「そうか、何人中何人位がスーパーグリーンを買いたいと考えるかですね」

土屋 「そうすると母集団にあたるもの、つまり視聴率でいう全国の家庭は何になる？」

佐藤 「全国の消費者ですかね」

土屋 「全国のことを知りたいのか」

佐藤 「MCスーパーについてです」

土屋 「日本中からMCスーパーに来るのか？紅白の視聴率にアメリカの家庭は入ってないぞ」

佐藤 「そうかMCスーパーに来る客だ」

土屋 「MCスーパーに来ない客はだめか？」

佐藤 「来ない客でも買いたいと思う人はいますよね」

土屋 「だったら母集団である全国の家庭にあたるものは何だ？全国の家庭でもテレビのない家は数に入れないぞ。つまりテレビを見ることができる、紅白を見ることが出来る日本の家庭だろう」

佐藤 「そうか。MCスーパーに来店する客というより、来ることが出来る客か。MCスーパーの商圈*内にいる顧客だ」



商圈

対象顧客の地理的範囲のこと。メーカー、卸、小売などで幅広く、さまざまな意味に使われる。小売業でいう商圈とは「店で販売される商品を購入する確率がゼロより大きくなる潜在的顧客を含む地域」と定義されるのが一般的。

土屋 「じゃ、標本は何だ？つまり視聴率調査機についている家庭にあたるものは何だ？」

佐藤 「そうか、MCスーパーにこの3カ月で来店した客ですね。でもそうになると昨日も今日も来た人は、同一人物でも、2人と数えるのですか？」

土屋 「だんだんわかってきたじゃないか。母集団はスーパーに来店することができる顧客で、標本は実際の来店客だ。かつ毎日の売上を見るなら、昨日来たAさんと今日来たAさんはちがう客とみなすわけだ」

佐藤 「4月7日のMCスーパーに来ることのできる顧客、4月8日に来ることのできる顧客・・・とやっていけばいいんですね」

土屋 「標本値は何だ？」

佐藤 「MCスーパー中央店の月報にある毎日の売上個数、これが標本値ですね。でも商品が4種類で13週91日分あるから、 91×4 で標本値が多すぎますよね」

土屋 「だから統計量に加工するんだらう。視聴率にあたるものだ。それぞれの家庭が見たか見ないかのデータを持ってるんでは大変だらう」

佐藤「ということは平均値だな。平均日販*ってやつですよ。これなら計算してある。えーと、スーパーグリーンが1日9個。ライバルのナショナルヘルスハイパーが1日13個か。えー1.4倍も向こうの方が上だ」

平均日販

毎日売れる商品（最寄品という）は1日当たりの売上個数、金額が最も大切な指数となる。これを平均日販、略して日販と言う。

土屋 「佐藤が知りたいのは4月から6月の3ヶ月間の結果か？それが母集団か？」

(図1-6)

佐藤 「もう終わったことが知りたいことではないですよ。終わったことから考えて、明日はどうなるかを知りたいんです。そうすると知りたいことは、明日以降のことなので母集団は明日以降の消費者だ。MCスーパーの販売計画から考えて向こう3ヵ月位かな。でもそれでは標本もないし、標本値もないので、平均日販も計算できないなあ」

土屋 「昨日と今日の消費者はまるでちがう買物をするのかな？」

佐藤 「昨日と今日で全然違うということはないですが、まるで同じということもありません。全然違うなら、昨日までの結果である日販なんて見てもしようがありませんよね。でも、毎日毎日、スーパーグリーンはテレビコマーシャルを流し、販売促進をやり、消費者がそれを飲み、と状況は変化しているんですから、まるで同じ買物をするということはないですよ」

土屋 「だったら2つに分けて考えよう。同じ部分と違う部分だ。同じと考えたときの母集団は何だ？」

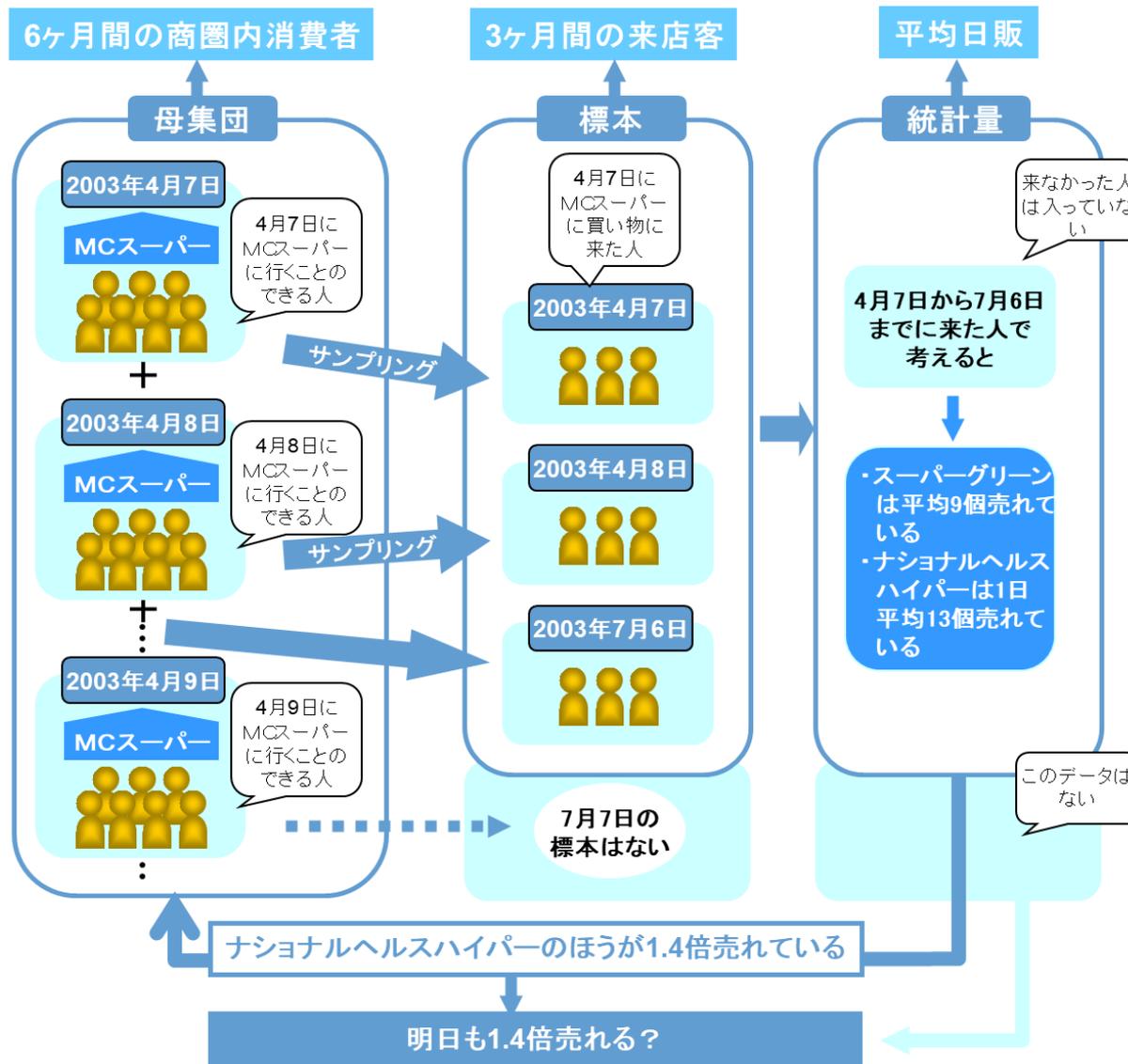


図1-6

佐藤 「過去3ヶ月と向こう3ヵ月間を同じように買物をすると考えるのですから、MCスーパーに来店することのできる6ヵ月間の消費者ですよ。ここから過去3か月分の来店客が標本としてサンプリングされ、日販という統計量が計算されるわけだ。つまり同じ買い物をすると考えるとナショナルヘルスの方が1.4倍売れるわけだ」

土屋 「母集団には『わかっている数字』と『わかっていない数字』があって、『わかっている数字』を使って『わかっていない数字』を予測していくわけだ」

佐藤 「6ヵ月間の商圏内消費者という母集団のうち、『過去3ヵ月間に実際に来店した顧客がスーパーグリーンやナショナルヘルスハイパーを買った個数』が、わかっている数字ですよ」

土屋 「そのとおり。この『わかっている数字』が標本であり、それを加工した統計量だ」

佐藤 「『わかっている数字』は、この商圈内消費者が次の3ヶ月でスーパーグリーンやナショナルヘルスハイパーをどれくらい買うか、ですよ。『わかっている数字』で『わかっていない数字』をあてるわけですか。でもそれだとやはり、1.4倍向こうが売れると考えるしかないでしょう」

土屋 「それなら佐藤の出番はないぞ。そんな単純な計算だけで予測するんなら人間なんていらんよ。過去と将来は『違い部分』もあるとさっき自分で言ったろう。『同じ部分』はわかっている数字で平均日販だ。それじゃあ『違う部分』のわかっている数字は何だと思う？」

佐藤 「過去3ヶ月と将来3ヶ月の『違う部分』で、わかっている数字なんてないでしょう？」

土屋 「過去3ヶ月は13週、91日のデータから成り立っているだろ？4月1週と4月2週の平均日販はわかっているよな。同じか」

佐藤 「それは違いますよ」

土屋 「13週、91日のデータは皆違ってはいるけど、何かその違いに『傾向』のようなものがないか」

佐藤 「そうか、この3ヶ月のデータに傾向のようなものがあれば、次の3ヶ月もこの傾向に近い『違い』を出すと考えればよいのか。傾向っていうくらいだから、増加傾向とか減少傾向ということですよ。ちょっと考えてみます」

(図1-7)

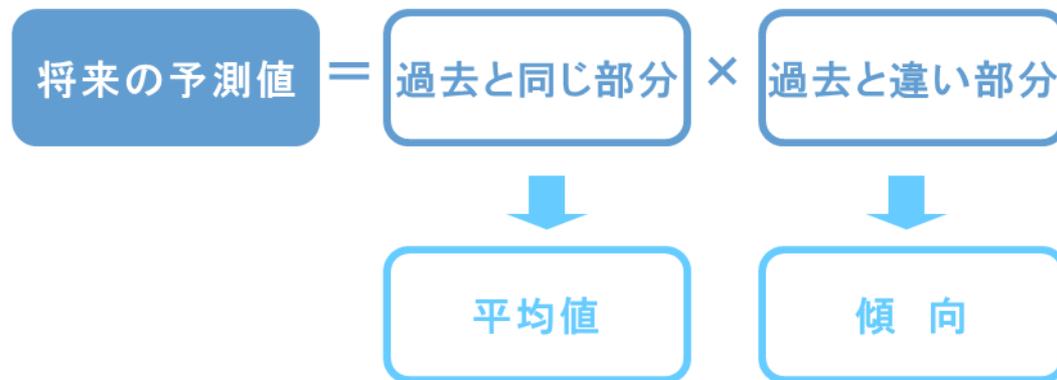
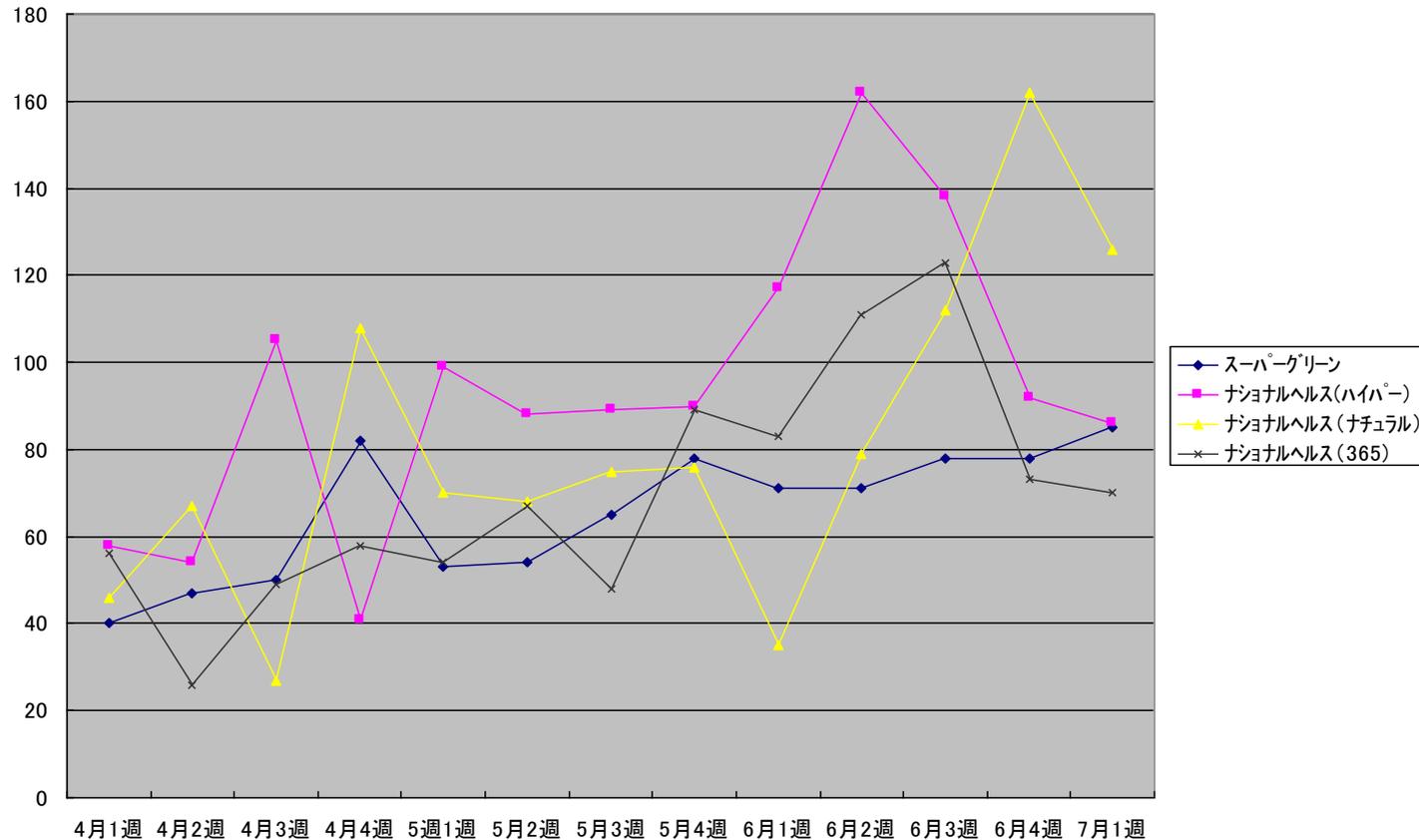


図1-7

佐藤は先ほどの表計算ソフトのグラフをもう一度見てみた。(図1-3)



佐藤「このグラフを見ると、どうもスーパーグリーンは時とともに増加しているように見えるぞ。ナショナルヘルスハイパーは、増えたり減ったりでわかりづらいけど、増えているように見えるなあ。元の数字で見てみよう。そうか、こういう折れ線グラフは傾向を大まかにとらえるものなのか。そういえばこのグラフを見せた時、支店長に『これをMCスーパーに持って行って何というつもりだ』と言われたなあ。『これは過去と違う部分を探すために傾向をざっと見るのに使います』と答えれば良かったのか」

佐藤は1週目から12週目までの日販を追いかけていった。そして最後の13週目まで来た。(図1-8)

第13週

		6/30(月)	7/1(火)	7/2(水)	7/3(木)	7/4(金)	7/5(土)	7/6(日)	週合計	週平均
スーパーグリーン	個数	9	8	10	14	16	14	14	85	12.1
	売上額	2250	2000	2500	3500	4000	3500	3500	21250	3035.7
ナショナルヘルス (ハイパー)	個数	20	3	19	12	4	25	3	86	12.3
	売上額	6000	900	5700	3600	1200	7500	900	25800	3685.7
ナショナルヘルス (ナチュラル)	個数	20	18	16	4	32	6	30	126	18.0
	売上額	6000	5400	4800	1200	9600	1800	9000	37800	5400.0
ナショナルヘルス (365)	個数	17	2	14	6	22	4	5	70	10.0
	売上額	2550	300	2100	900	3300	600	750	10500	1500.0

佐藤 「スーパーグリーンは4月1週では平均日販数は6個で、13週目には12個になっているぞ。何だ倍じゃないか。ナショナルヘルスハイパーは4月1週では8個で、13週目は12個か。ハイパーも1.5倍で増えているけど13週目にはスーパーグリーンが追いついてるぞ。・・・そうか、4月から6月でスーパーグリーンが平均13個でも、過去増加傾向にあるなら、明日も昨日より増加すると考える方が普通だよな。よく考えればこの増加傾向だって商品力だ」

佐藤は土屋の所へかけつけた。

佐藤「土屋さんわかりました。商品力には『昨日、何個売れたか』という平均日販の他にも『昨日と比べてどのくらい伸びているか』という増加率のようなものがあると思います。スーパーグリーンは明らかにナショナルヘルスハイパーより率が高いです。スーパーグリーンはナショナルヘルスハイパーに4月には負けていましたが、6月には追いついています。今後はきっと逆転していく力があると思います。これも商品力ですよね」

土屋「スーパーグリーンはなぜ増加率が高いんだ」

佐藤「それも数字でわかるんですか」

土屋「わかるわけないだろう。数字は結果だ。結果から人間が予測するんだ。これが仮説だ」

佐藤「そういえばうちの社長がいつも『営業は仮説の検証だ』と言ってますよね。その仮説を教えてください」

土屋「俺よりも君が考えるべきだろう。君の方がスーパーグリーンもMCスーパーもよく知っているはずだ」

佐藤 「MCスーパーの売場は健康飲料についてはこの3ヵ月でフェースは変えてないし、POP広告*も打ってないので、それが増加原因ということはないですよ。スーパーグリーンもナショナルヘルスも商品的に変わってないし……。そうか消費者、MCスーパーの顧客が変わったんだ」

土屋 「4月と6月の消費者はなぜ変わったんだ」

佐藤 「テレビコマーシャルを毎日打っているから、スーパーグリーンを知らなかった人が知るようになったんだ。そして、スーパーグリーンを一度買ってくれた人が、その後、浮気せずに何度も買ってくれているんだ。そういえば前に受けたマーケティングセミナーで、テレビコマーシャルは認知広告*としては最高のメディアだと言ってたっけ。『健康食品は他の商品と比べて一度買った商品を続けて買う確率、つまりリピート率が高い』とも言っていた。そうか、増加率は認知度アップとリピート率から成り立っているんだ」

([図1-9](#))

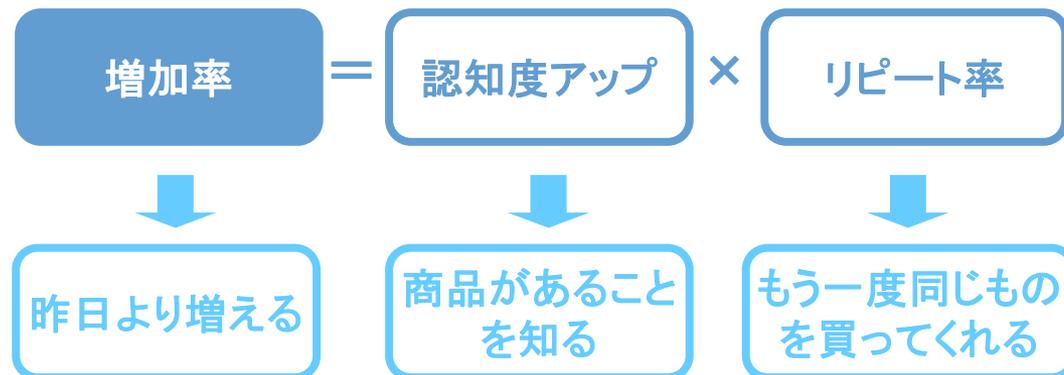


図1-9

POP広告

POPとはPoint Of Purchaseの略。購買時点広告のことで、店舗内での商品広告をさす。

認知広告

広告は大きく、製品広告（製品を宣伝）と企業広告（企業を宣伝）に分かれる。製品広告は認知広告、機能・品質広告、ブランド広告、価格広告などのテーマにより分けられる。認知広告とは新製品などを消費者に知ってもらうために打つ広告を言う。

佐藤 「リピート率はまさに商品力ですけど、認知度のアップを商品力と言うのは少し変かもしれませんね。認知度アップは、平均日販、リピート率という商品力を『本当の数字』にするためのものですよね。世の中に商品を知らしめて、商品力を発揮させる環境を作ってるんですよ」

土屋 「商品力とは消費者がその商品を買う率だったろう。確かにスーパーグリーンを知らない人が買う確率なんて意味がないよな。きちんとスーパーグリーンを知ってもらって、買うか買わないか決めてもらいたいものね」

佐藤 「よしまずは増加率を調べてみよう。それからリピート率だ」

佐藤「増加率ということは伸び率だな。日々の増加率では数字が多すぎてあまりよくわからないから、週ごとにやってみるか」

佐藤は表計算ソフトを使って、対前週の増加率を出してみた。

対前週増加率

	第2週	第3週	第4週	第5週	第6週	第7週	第8週
スーパーグリーン	1.18	1.06	1.64	0.65	1.02	1.20	1.20
ナショナルヘルス(ハイパー)	0.93	1.94	0.39	2.41	0.89	1.01	1.01
ナショナルヘルス(ナチュラル)	1.43	0.42	3.86	0.65	0.97	1.10	1.01
ナショナルヘルス(365)	0.46	1.88	1.18	0.93	1.24	0.72	1.85

第9週	第10週	第11週	第12週	第13週	平均
0.91	1.00	1.10	1.00	1.09	1.09
1.30	1.38	0.85	0.67	0.93	1.14
0.47	2.19	1.42	1.45	0.78	1.31
0.93	1.34	1.11	0.59	0.96	1.10

図1-10

佐藤「土屋さん、平均増加率を出してみました。スーパーグリーンが1.09でナショナルヘルスハイパーが1.14です。あれ、うちが負けてる。何か変だなあ」

土屋 「増加率の平均ってどうやって出した？」

佐藤 「えっ、増加率の12個の数字を足して12で割ったんですが・・・」

土屋 「それじゃダメなんだ。例えば毎年10、20、80と増えているとき、増加率は10から20で2倍、20から80で4倍だろ。平均増加率は？」

佐藤 「2倍と4倍だから3倍ですか」

土屋 「年々平均して3倍増えているのか。ということは、2年で9倍になるはずだろう。10から80の8倍にしかなっていないじゃないか」

佐藤 「あれ？」

土屋 「そういう場合は幾何平均というのを使うんだ。手元の本で調べてみろ」

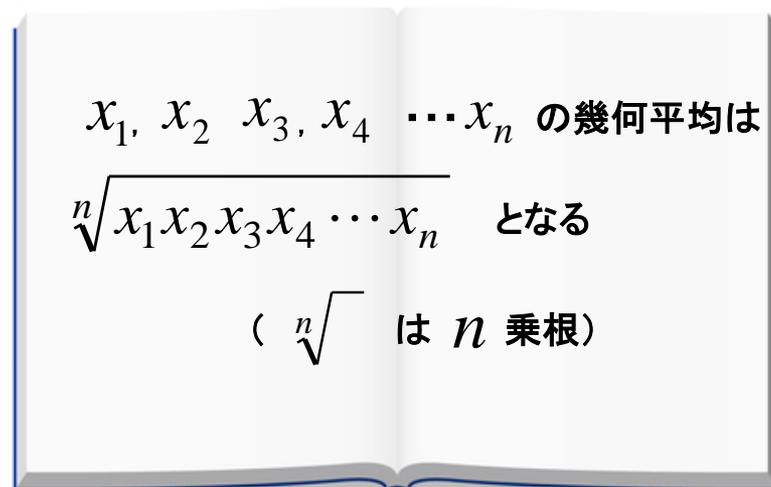


図1-11

佐藤 「 n 乗根って何ですか」

土屋 「 2×2 は4だろう。このとき4の2乗根は2だ。同じ数字で何を2回かければ4になるかだ。2乗根は普通は平方根といい、ルートという記号を使っている。3乗根は何を3回かけるかだ。さっきの対前週の平均増加率を出すには、同じ数字で何を12回かけると、この増加率を12個かけ合わせたものになるかを出せばいい」

佐藤 「電卓にないんですけど」

土屋 「表計算ソフトでやるんだ」

対前週増加率(幾何平均)

	第2週	第3週	第4週	第5週	第6週	第7週	第8週
スーパーグリーン	1.18	1.06	1.64	0.65	1.02	1.20	1.20
ナショナルヘルス(ハイパー)	0.93	1.94	0.39	2.41	0.89	1.01	1.01
ナショナルヘルス(ナチュラル)	1.43	0.42	3.86	0.65	0.97	1.10	1.01
ナショナルヘルス(365)	0.46	1.88	1.18	0.93	1.24	0.72	1.85

第9週	第10週	第11週	第12週	第13週	幾何平均
0.91	1.00	1.10	1.00	1.09	1.07
1.30	1.38	0.85	0.67	0.93	1.03
0.47	2.19	1.42	1.45	0.78	1.09
0.93	1.34	1.11	0.59	0.96	1.02

図1-12

佐藤 「スーパーグリーンは1.07で、ナショナルヘルスハイパーは1.03か。スーパーグリーンは週に平均して7%伸び、ナショナルヘルスハイパーは3%か。ということは倍以上か。よし勝っているぞ。あれ、ナショナルヘルスのナチュラルは9%か、もっと大きいや。でもナチュラルは血糖値を意識しているから、うちのスーパーグリーンと必ずしも競合しているとは言えない。増加率は認知度のアップ分とリピート率の合わさったものだ。土屋さん、何とかこのリピート率だけをもう少しはっきり数字で出せないないですか。増加率だけではMCスーパーへのアピール度が弱いですよ。はっきりとリピート率もスーパーグリーンが勝っていると言いたいです」

土屋 「そういうときは、同じようなことを昔考えた人がいると思えばいい」

佐藤 「こんなことで悩んでいる人はたくさんいますよね」

土屋 「そのうちの賢い人の考え方を利用する、それが学問だ」

佐藤 「じゃ、教えてください」

土屋 「まずは自分で考えてみなきゃ。何を知りたいのかもっとよく考えて。それが見えてきたら、賢い人の考え方を教えてあげる」

佐藤 「そんな冷たいことを言わずにヒントだけでも」

土屋 「君たちは本を読んでヒントになることを見つけようとするからだめなんだ。まず何を解決しなくてはならないかを自分で考え、それが見えてきたら、その方法を調べていかなきゃ」

佐藤 「さっきのグラフを一生懸命見てるんですが、何も浮かびません」

土屋 「そのグラフは週ごとの売れ行きを見るために作ったんだ。さっきのように増加・減少などの傾向を見るには最適なグラフだ。しかし、逆に言えばそれ以外の目的には使えないんだ。元の数字を見て、自分の頭を使って、何か、“気づき”のようなものを探せ」

佐藤はMCスーパーからもらった364個の数字をじっと見ていた

「やっぱり土、日はどれもよく売れているなあ。あれ、ナショナルヘルスハイパーは4月の土曜日がすごく売れているぞ。やっぱり土曜日が最高かあ。でも5月の土曜日はあまり売れていない……。スーパーグリーンはどうかな。4月の土曜日は随分負けているなあ。でも月曜日は勝っているときもある。待てよ、こうしてみるとうちの商品はコンスタントに売れていて、ナショナルヘルスは波が大きいなあ。コンスタントに売れるということは、スーパーグリーンに固定客がついて、定期的に買ってくれていると考えられるよな。そうか、これがリピート率だ。リピート率はコンスタントに売れているかどうかを見ればいいんだ」

翌日、佐藤は土屋にこのことを報告した。

佐藤「リピート率をどう考えればいいのかわかりました。コンスタントに売れているものがリピート率が高いといえます。そして、スーパーグリーンの方がナショナルヘルスよりコンスタントに売れているので、きっとリピート率が高いと考えられます」

土屋 「いいところに気づいたじゃないか、スーパーグリーンの方がどれくらいコンスタントなんだ？」

佐藤 「そう言われても『コンスタントさ』ってどうやって計算するんですか」

土屋 「そこまで考えれば、つまり『コンスタントさ』を出そうと自分で思えば、『コンスタントさ』の出し方を調べればいいんだ。いつもそう考えなきゃ。マニュアル読んで『コンスタントさを出せ』と書いてあるからといって、その通りに、意味もよくわからずに出すんじゃないぞ。数字を扱うときはデータの処理のやり方をまず勉強するのではなく、自分のカンを生かし、何を知りたいかを考えて、そのやり方を調べるんだ。まあ、そこまで考えたんだから教えてあげよう。『コンスタントさ』については、人間が長い間悩んだ末に結論を出したんだ。それが標準偏差という考え方だ。照準偏差はバラツキを表すもので『コンスタントさ』の反対だ。標準偏差が小さければバラツキは小さい、つまりコンスタントということだ」

佐藤 「でも、バラツキって何ですか。ちょっと本を調べてみます」

『誰でもわかる統計の本』には次のように書いてあった。

バラツキとは標本値が平均の中心からどのくらい離れる傾向にあるかを示すものであり、標準偏差(σ :シグマ)がよく用いられる。標準偏差は分散(σ^2)の平方根であり

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} \quad \text{で計算できる。}$$

\bar{x} は $x_1 \sim x_n$ の平均値

図1-13

佐藤 「大体この本は文章も難しいけど、数式を見ると、もう『まいった』と思います。わかりやすく教えて下さい」

土屋 「こういうとき大切なことは数式を覚えることじゃない。数式の意味を直感でとらえていくことだ。人間が意味を理解すれば、あとは数式をコンピュータが覚えていてくれて、早く正確に計算してくれる。いくら計算できてもその数式にどういう意味があるのかわからなければ、その結果を使えないだろう」

佐藤 「じゃあ、標準偏差の意味を教えてください」

土屋 「平均値はわかるだろ。標準偏差っていうのはこの平均値からどれくらい離れているかを平均するものだ」

佐藤 「『平均値からどれくらい離れている』ってどういうことですか」

土屋 「平均近くに数字が集まっていることをコンスタントっていうんだ。よく『うちの社員は粒ぞろいだ』と言うだろう。これがコンスタントだ」

佐藤 「うちのセールスは成績に関しては『どんぐりの背比べ』です。これもコンスタントか」

土屋 「じゃあ、やり方を説明してあげよう。例えばうちのスーパーグリーンが9個、11個、8個、12個売れたとする。平均は何個？」

佐藤 「えーと、10個です」

土屋 「ナショナルヘルスが6個、14個、5個、15個売れたとすると平均は何個？」

佐藤 「えーと、やっぱり10個です」

土屋 「ナショナルヘルスの社員だったら、この最大の15個が商品力というはずだ。5個、6個のときは売り方が悪い。本来なら15個まで売れる力があると言うだろう。15個にするためにPOP広告を打ちましようと言うよな。じゃあ、うちの商品力は12個か？それとも平均10個でナショナルヘルスと同じなのか」

佐藤 「どう見ても、うちのスーパーグリーンの方がコンスタントに売れていて、店にとってはよさそうですね」

佐藤 「えーと、4つを足すと、あれ、0です」

土屋 「そのコンスタントさを計算するのが標準偏差だ。スーパーグリーンは9個、11個、8個、12個と売れて、平均10個だ。最初の9個は平均10個から『1個』離れているだろう。これをもう少し正確にして『その数から平均を引く』と決めると、9個のときは9引く10で-1となる。これが偏差だ。差とは引き算のことで、偏はかたよりだ。偏差の平均が標準偏差だ。スーパーグリーンの偏差は-1、1、-2、2となる。平均してみろ」

土屋 「そうなんだ。こうするとプラスとマイナスが打ち消し合って偏差の平均はいつも0になってしまう。よく考えればあたりまえで、だから平均なんだ。こういうときは数学には奥の手があるんだ。偏差を2乗してしまうんだ。こうすると全部プラスの数字になる。そのうえで、その平均をとる。これを分散というんだ。そしてさっき2乗したから、これを戻すためにその平方根をとる。これが標準偏差で、さっきの本に書いてある式だ」

佐藤 「うーん、まいった。わかるようでわからない」

土屋 「難しく考えなくてもいい。式はコンピュータが覚えてくれるんだから。君は標準偏差は『偏差の平均』と理解すればOKだ。これで計算すると、スーパーグリーンの9個、11個、8個、12個では標準偏差は1.6だ。ナショナルヘルスの6個、14個、5個、15個では4.5だ。標準偏差つまりバラツキは約3倍だ。つまり、スーパーグリーンの方が3倍コンスタントだ。どうだ、数字を見たときの自分の直感と合うか？」

佐藤 「いい感じですよ」

土屋 「標準偏差の出し方も昔は色々意見はあったらしいが、なるべく人間の直感に合うように試行錯誤したんだ。まあ『9、11、8、12』は『6、14、5、15』に比べてバラツキが1/3、コンスタントさが3倍、ということなら納得いくところだろう。もらったデータで標準偏差を計算してみたら。データが表計算ソフトに入っているなら簡単にできるぞ」

佐藤は表計算ソフトを使って、スーパーグリーンとナショナルヘルスハイパーの標準偏差を出してみた。

佐藤 「うちのスーパーグリーンは3.9、ナショナルヘルスハイパーは8.4となりました。えっ、倍か」

土屋 「スーパーグリーンの方が2倍コンスタントさがあるわけだ。次に佐藤が考えるのはこのコンスタントさはお店にどういう影響を与えるかだ」

佐藤 「そう言えば先輩が前に、コンビニはコンスタントに売れる商品しか置かないから、成功したと言っていました。考えてみます」 ([図1-14](#))

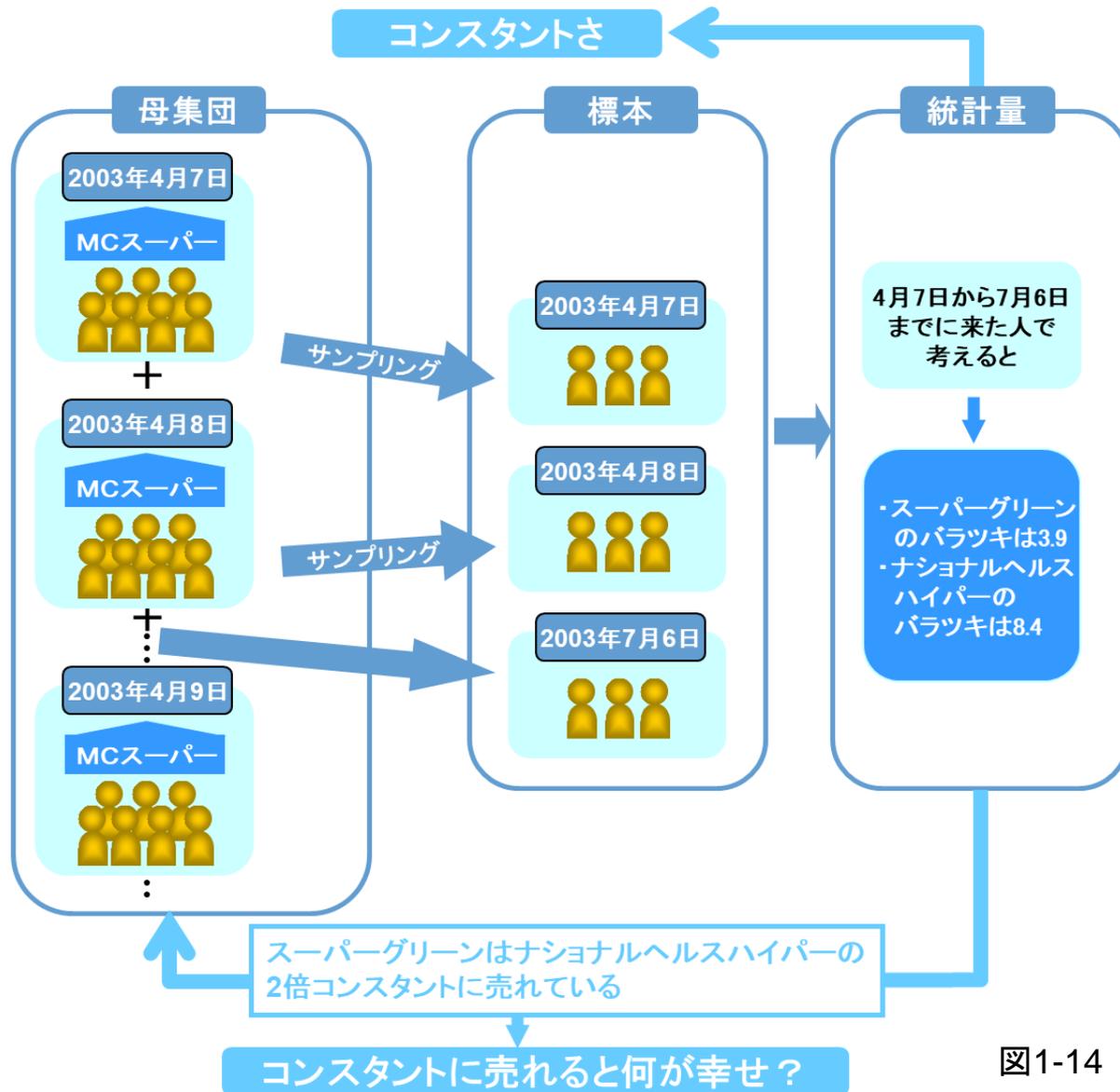


図1-14

佐藤はMCスーパーにデータの分析結果を報告しに来ていた。

佐藤「今回は商品カテゴリー*上、重複しているスーパーグリーンとナショナルヘルスハイパーの2つについて比較してみました。日販数は13週平均でスーパーグリーンが9.4、ナショナルヘルスハイパーが13.4となっております。また、対前週の売上増加率を見ますと、12週平均でスーパーグリーンは7%、ナショナルヘルスは3%であり、スーパーグリーンの増加率はナショナルヘルスハイパーの2倍以上になっています。次に91日の日販数の標準偏差を見ますと、スーパーグリーンが3.9、ナショナルヘルスが8.4となっております」

商品カテゴリー

生鮮食品などの商品部門と、アイテム（商品品種のこと。マグロの刺身など）の中間に位置する商品分類のこと（例えば刺身など）。店舗では取り扱う商品の分類を商品部門（大分類）、商品カテゴリー（中分類）、商品アイテム（小分類）と考えていくことが多い。単にカテゴリーとも言う。

木村 「標準偏差って何？」

佐藤 「バラツキを表すもので、平均から離れている度合いです。小さいほどその商品がコンスタントに売れたということを意味しています。同じ1日平均10個売れるのでも、9、11、8、12と売れるか、6、14、5、15と売れるかの違いです。つまり、うちのスーパーグリーンの方がナショナルヘルスハイパーより2倍以上コンスタントに売れています」

木村 「コンスタントに売れると何がいいの？」

佐藤 「その前に私の仮説を聞いてください。

健康飲料は毎日一定量飲むのが普通です。したがってコンスタントに売れていくのが当然といえば当然です。当社のスーパーグリーンは先ほど述べたように、日々、日販数が増加しております。また、他店のデータを見ても同じ傾向を示しています。逆に言えばコンスタントさを失って、平均から離れていることとなります。それでもスーパーグリーンの標準偏差がナショナルヘルスハイパーより小さいこと、つまりコンスタントであることを見ると、購入客が着実に増え、なおかつその顧客にコンスタントに購入していただいていると考えられます」

木村 「ふむ」

佐藤 「購入者が増えているのは、テレビコマーシャルなどの認知広告が効いているためだと思います。一方、ナショナルヘルスは増加率が小さく、逆に標準偏差が大きいことを考えると、とりあえず飲んでみる顧客は多いが、それが続かないのだと考えられます。健康飲料のように目的買い*の商品がコンスタントさを保つということは、リピート率が高いということであり、お客様が常に補充買い*をしてくれることです。ということは、MCスーパーさんにとってもいわゆる固定客の獲得となる効果を持つと考えられます。そのためにもテレビコマーシャル同様にスーパーグリーン認知度を高めるべく、フェースの増加や定期的な製品紹介のためのPOP広告などが効果的と考えます」



目的買い

店舗内で買う商品、または買うこと自体を決めることを衝動買いという。一方で店舗に入る前から商品購入を決めているものを計画買いという。計画買いの中で、店舗へ来た目的がその商品を購入することにあるものを目的買いと言う。スーパーグリーンを買いだいたいと思ってMCスーパーに来店すれば目的買いとなる。

補充買い

消費者が常に一定量消費している商品で、なくなったり、少なくなると定期的に購買するもの。

佐藤 「お手元の最終の13週のデータを見てください。この週は、両者の週販はほとんど同じです。しかしスーパーグリーンは1日最大で16個、ナショナルヘルスハイパーは最大25個売られています。両者を欠品が起こることのないように陳列するとしたら何個必要でしょうか、スーパーグリーンは16個、ナショナルヘルスハイパーは25個です。同じ平均日販12個を得るのにどちらが得かを考えてみて下さい」

第13週

	6/30(月)	7/1 (火)	7/2(水)	7/3(木)	7/4(金)	7/5(土)	7/6(日)	週合計	週平均
スーパーグリーン	9	8	10	14	16	14	14	85	12.1
ナショナルヘルス (ハイパー)	20	3	19	12	4	25	3	86	12.3

図1-15

木村 「これって3ヵ月のデータで大丈夫なの？」

佐藤 「母集団、つまり本当の値を知るにはもっとデータが必要です。できれば引き続きデータを分析させてください。ただ分析するだけではもったいないので、先ほどの仮説を立証する意味でも店舗にスーパーグリーンのフェースを増やし、POP広告を打つことで来店客数、日販数の変化を見てみたいのですが、いかがでしょうか」

期待値

MCスーパーでくじによるキャッシュバック・キャンペーンを始めました。5000円以上買った客に箱の中からボールを取り出してもらい、その色で1等600円、2等300円、3等100円、4等0円(はずれ)として現金を返すというものです。MCスーパーは1~4等のボールを箱に同数だけ入れ、誰かが1つボールを取り出したら、ボールを箱に戻すという形で行いました。箱の中にボールが入っている割合は客にはわかりません。

何人かがボールを取り出したところ、次のような結果になりました。

0、100、0、300、100、0、300、100、0、300

これを見ていた客が平均値を計算しました。120円です。これをこの懸賞に期待値と考えてよいでしょうか。期待値とはまさにボールを1回引くとき「期待できる値」です。母集団である箱の中にはボールが同数入っているのですから、それぞれのボールを引く確率は1/4です。したがって、母集団の本当の期待値は

$$600 \times \frac{1}{4} + 300 \times \frac{1}{4} + 100 \times \frac{1}{4} + 0 \times \frac{1}{4} = 250$$

「250円」となります。しかし客はこの確率(ボールが入っている割合)がわからないので、標本の統計量である平均値「120円」で母集団の期待値を予想するしかありません。これが数学でいう「推定」です。

気づいた人もいるかもしれませんが、このくじを延々と引き続けると、しだいにこの標本の平均値は期待値250円に近づいていきます。これが「大数の法則」です。

あるものを調査するとき、一般には母集団すべてのデータを調べるのは不可能です。サンプリングした標本の統計量を見るしかありません。このとき、標本量が増えるほど、母集団の値に近づいていくこととなります。しかし一方で、サンプリングして調査するコストが上がってしまいます。このトレードオフの関係を考えて適当な量の標本にすることが統計調査の基本といえます。

第 2 章

明日はいくつ
売れるかを予測

ミドリ食品では取引先の小売店舗との関係を密にし、両者がパートナーシップを組んで販売していく方法を模索していた。いわゆるSCM*の実現である。その第一歩として小売店との間で毎月できれば年間の、ミドリ食品の商品の店舗における販売目標を共有し、これを達成するために両者が努力するという方向に出来ないかと考えていた。

中村支店長「佐藤、MCスーパーとはうまくいっているか？」

佐藤「スーパーグリーンのデータ分析以来、いろいろ頼りにされるようになりました」

中村「そうか。ところで君も当社がSCMにチャレンジしようとしているのは聞いたことがあるだろう」

SCM

Supply Chain Managementの略。消費者、利用者に対して小売・卸・製品メーカー・資材メーカーなどが1つの企業体のようになって活動していくこと。4ここではMCスーパーとミドリ食品が提携し、共同で消費者へタイムリーな商品提供を考えていくこと。

佐藤 「はい。でも何をやりたいのかよくわかりません」

中村 「要するに君たち営業マンが小売店などの顧客のパートナーとなって、店舗の売上増大のために働き、それが結果として当社に利益をもたらすと考えることだ。MCスーパーとの関係が良いなら、そこで試しにやってみるか」

佐藤 「何をしたら良いのですか？」

中村 「とりあえずMCスーパーのどこかの店で、スーパーグリーンの販売目標を一緒に作り、それを達成したら販売支援金*として当社から報奨金を出すようにしよう」

佐藤 「どうやって目標を作ったらよいのですか？」

中村 「それを考えろ。MCスーパーはなるべくバーを低くしたいし、うちとしてはなるべくバーを高くしたい。そこを調整するんだ。ちょっと難しいかもしれないけど、がんばってくれ」

販売支援金

メーカーなどが取引先の起業に対して販売促進の目的で支払うものをリベートという。リベートは販売支援金、報奨金、販売促進費とも呼ばれる。

佐藤はMCスーパーの工藤店長を訪ね、その趣旨を説明した。

工藤「おもしろそうだね。やってみようよ。とりあえずスーパーグリーンでやろうか。細かいことはバイヤーの木村とよく相談して決めてよ」

木村バイヤーも趣旨を理解してくれた。

木村「やるならちゃんとやりましょう。これがうちで店頭化*してからのスーパーグリーンの月販数。お宅が『300個にしましょう』とって、うちが『そんなの無理だ。250個でしょう』と交渉していくのは何か声の大きい方が勝つみたいでいやだよ。しかもそれじゃ長続きしないよね。互いが納得できる数字にしよう。私の感じとしては、『普通にやれば売れる数字』というので決めたいよな」

手許にある38個の数字を見つめながら、佐藤はこの「普通にやれば売れる数字」というのが頭から離れなくなってしまった。(図2-1)

店頭化

新商品などが、特定の店舗に初めて陳列されること。

スーパーグリーン月別販売個数

	2000年	2001年	2002年	2003年
1月	—	224	231	346
2月	—	206	278	298
3月	—	188	229	307
4月	—	201	266	358
5月	120	198	225	388
6月	126	246	235	372
7月	142	233	282	—
8月	130	211	306	—
9月	128	240	264	—
10月	146	222	312	—
11月	205	213	288	—
12月	182	255	319	—

図2-1

佐藤はもらったデータを表計算ソフトに入れ、コンピュータと格闘していた。

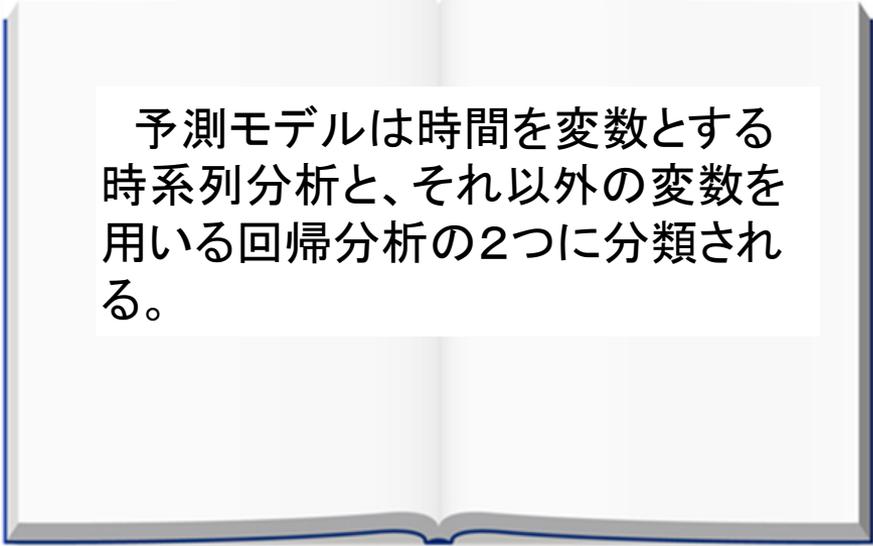
「とりあえず、この38個のデータの平均を出してみよう。240個か、意外と少ないなあ。2003年に入ってからは大体300個以上売れているのに何かしっくりこないなあ。しかし平均値以外どうしたらいいんだ。標準偏差を出しても意味がなさそうだし…」

佐藤は「誰でもわかる需要予測」という本を買い、読んでみた。しかしやたらと難しい式が並んでおり、それにやり方もたくさんあって、どれを使ってよいかもわからなかった。

「そうだ、考え方だ。式よりも考え方を見よう。最初の方に書いてあるはずだ」

その本の最初に書いてあったのは図2-2のようなことだった。

「変数とは数字が変わるということか。今やろうとしているのは売上と時間のデータしかないから時系列分析だ。ここを読めば良いのか

An illustration of an open book with a white text box on the left page. The text box contains the following text:

予測モデルは時間を変数とする
時系列分析と、それ以外の変数を用いる
回帰分析の2つに分類される。

図2-2

時系列分析の項目を見るとまず図2-3のようなことが書いてあった。

「時系列データというのはこの場合、月販数だ。TCSIとは要するにこの4つに分けて考えるということか」

「傾向変動はトレンドか。前にやった増加率のようなものだな。何となくわかるぞ。このデータはどう見ても右肩上がりで増加傾向だ」

「循環変動はサイクルか。上がると下がるをくり返していくやつか。景気みたいなもんだな。このデータは上がると下がるを定期的にくり返しているとはいえないなあ。ということは循環変動はなしか」

「季節変動はシーズン特性みたいなもんだな。水着は夏に売れ、冬はダメというようなものか。あるかもしれないな。よし月別の平均値を出してみよう」

(図2-4)

TCSIモデル

時系列データは次の4つによって構成されていると考える

- ・傾向変動 (Trend) …
 長期的な動きを示すもの
- ・循環変動 (Circulation) …
 一定期間内に増加と減少を繰り返すもの
- ・季節変動 (Season) …
 1年を周期として季節ごとの特性を示すもの
- ・不規則変動 (Irregular) …
 上記どの3つにもあてはまらないもので、予測不可能の誤差とも考えられる。この不規則変動以外のものを予測していくのが需要予測モデルである。

図2-3

スーパーグリーン月別販売個数(平均)

	2000年	2001年	2002年	2003年	平均
1月	—	224	231	346	267
2月	—	206	278	298	261
3月	—	188	229	307	241
4月	—	201	266	358	275
5月	120	198	225	388	233
6月	126	246	235	372	245
7月	142	233	282	—	219
8月	130	211	306	—	216
9月	128	240	264	—	211
10月	146	222	312	—	227
11月	205	213	288	—	235
12月	182	255	319	—	252

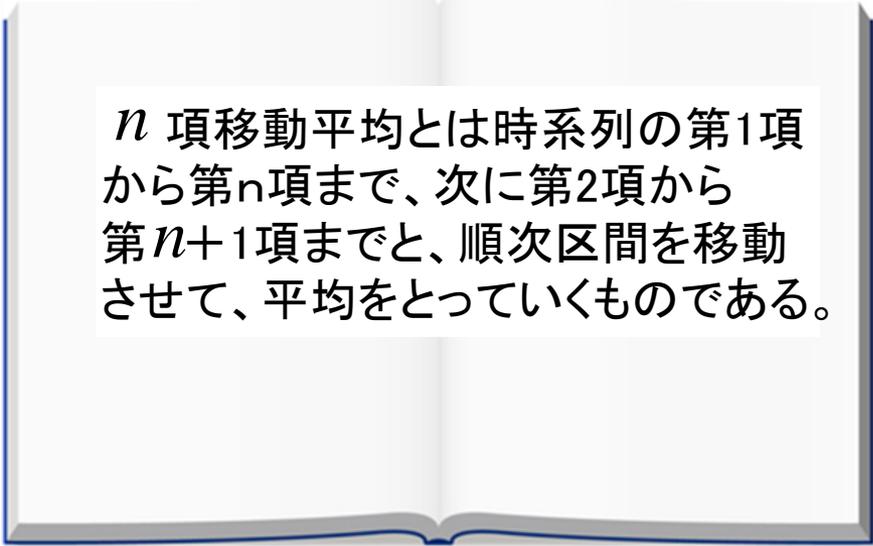
図2-4

「あれ、7月から随分落ちるなあ、そうか、増えている2003年のデータがないからか。2001年と2002年のデータだけで平均してみると1月から12月に向かって増えていってるなあ。まあ右肩上がりだからあたりまえか。これは季節のせいではないな

「不規則変動は？そうか、前に土屋さんが言ってた、昨日と明日はちがうというやつか。でもそれはこの本によると予測できないものという意味だな。そうすると傾向変動だけを考えれば良いのか。よし傾向変動を勉強しよう」

そこで傾向変動の所を読んでみると難しい数式がたくさん並んでおり、見たこともない文字まであった。

「まいったな、チンプンカンプンだぞ。あれ、この移動平均って易しそうだな。これにするか」(図2-5)

An illustration of an open book with a white text box on the right page. The text box contains the definition of a moving average.

n 項移動平均とは時系列の第1項から第 n 項まで、次に第2項から第 $n+1$ 項までと、順次区間を移動させて、平均をとっていくものである。

図2-5

そこで傾向変動の所を読んでみると難しい数式がたくさん並んでおり、見たこともない文字まであった。

「この本の例を見ると3つずつ平均とってるな。僕も3つでやろう。2003年7月を予測するには2003年4月、5月、6月の平均をとれば良いのか。簡単だな。373個か。なかなか良い数字かもしれないなあ。『普通にやれば売れる数字』にフィットしそうだなあ。でもこれじゃ木村バイヤーからこんなにたくさんの数字をもらった意味がないぞ。3つしか使っていないものなあ。・・・6ヶ月にしてみよう。6ヶ月移動平均では345個か。減っちゃったなあ。あたりまえか。スーパーグリーンはだんだん伸びているものなあ。しかも平均の対象である月が多いほうが良いなら、全部の平均の方が良いことになって、前に出した単純な平均となってしまうぞ。・・・2ヶ月移動平均だと、えーと380個か。何か3ヶ月の時の373個よりありえそうな気がするぞ。しかしそれなら直前の6月の372個が予測値として一番良いということになってしまう。いくらなんでも当月が常に前月と同じと考えるのも変だ。うーんよくわからなくなってきた」

佐藤 「土屋さん、ちょっと教えてください」

土屋 「今はだめだ。会議の資料を急いでまとめなくてはならないんだ」

佐藤 「そんなこと言わず10分だけ下さい」

土屋 「何を知りたいんだ？」

佐藤 「この38個の数字から需要予測したいんです」

土屋 「いつの予測？」

佐藤 「とりあえず2003年7月の予測です」

土屋 「その『とりあえず』が気になるけど、何かやってみたの？」

佐藤 「移動平均を出してみたんですが、よくわからなくなってしまう。何だか直前の6月のデータを使うほうが良いとも思えるんです」

とって試行錯誤してきたことを説明した。

土屋 「直前のデータを使ったんじゃ不規則変動を考慮してないだろう。平均というのは不規則変動、つまり『たまたま』というのを取り除いて、本当の姿を見ようとするものなんだ。今回のように時系列データを全部使って、かつ直前を重視して不規則変動を取り除きたいなら、指数平滑法を使うんだな。これも平均のとり方の一種だ」

佐藤は自分の机に戻って考えた。

「そうか、平均というのは直前のデータだけじゃ、『たまたま売れた』というブレみみたいなものが入っているので、それを取り払ってトレンドを見ることなのか。不規則変動は『たまたま』で、それを取り除くのが平均か」

先ほどの本にも指数平滑法はのっており、図2-6のように書かれていた。

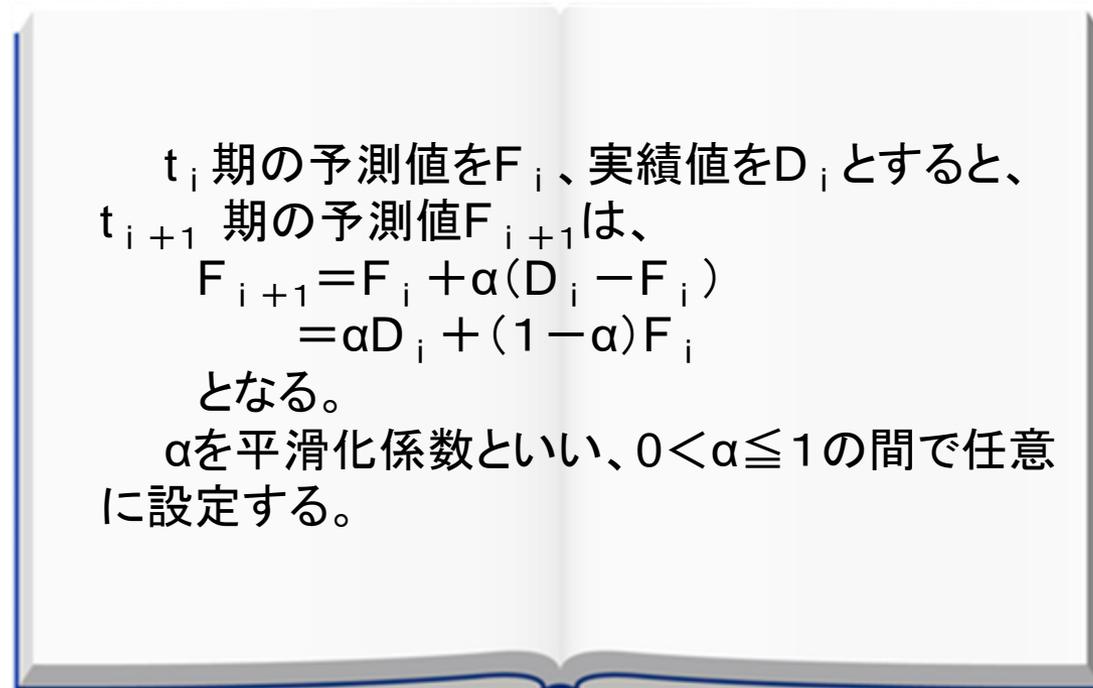


図2-6

「そうか、これは5期の予測値を4期の実績値と4期の予測値で計算するわけだ。よくわからないけど計算は簡単だ。これなら表計算ソフトでできそうだ」

「あれ、平滑化係数って何だろう。0から1の間といってもいくつにして良いかわからないなあ。本の例では0.3になっているからこれを使おうか。まず2000年5月を1期として2003年6月までを1列にする。これが本に書いてある実績値Dだ。隣に予測値Fの欄を作る。1期前のDに0.3、Fに0.7をかけるように表計算ソフトに式をセットする。あっ、1行目のF、1期目の予測値がないぞ。本では・・・と、1期の実績が入っている。そうしよう。・・・よしこれでできた。出たぞ。2003年7月は351だ。意外に低いなあ」

平滑化係数=0.3

		実績値(D)	予測値(F)
1期	2000年5月	120	120
2期	2000年6月	126	120
3期	2000年7月	142	122
4期	2000年8月	130	128
5期	2000年9月	128	129
6期	2000年10月	146	128
7期	2000年11月	205	134
8期	2000年12月	182	155
9期	2001年1月	224	163
10期	2001年2月	206	181
11期	2001年3月	188	189
12期	2001年4月	201	189
13期	2001年5月	198	192
14期	2001年6月	246	194
15期	2001年7月	233	210
16期	2001年8月	211	217
17期	2001年9月	240	215
18期	2001年10月	222	222
19期	2001年11月	213	222
20期	2001年12月	255	220

21期	2002年1月	231	230
22期	2002年2月	278	230
23期	2002年3月	229	245
24期	2002年4月	266	240
25期	2002年5月	225	248
26期	2002年6月	235	241
27期	2002年7月	282	239
28期	2002年8月	306	252
29期	2002年9月	264	268
30期	2002年10月	312	267
31期	2002年11月	288	280
32期	2002年12月	319	283
33期	2003年1月	346	294
34期	2003年2月	298	309
35期	2003年3月	307	306
36期	2003年4月	358	306
37期	2003年5月	388	322
38期	2003年6月	372	342
39期	2003年7月		351

これに
0.3かける
+これに
0.7かける

図2-7

翌日午後、佐藤はこの結果を土屋に報告した。

佐藤「土屋さん、できました。2003年7月は351個です。でも自分で何をやっているのかまるでわからないから、お客様にこの数字の意味を説明できません。指数平滑法って何をやっているんですか？」

土屋「2003年7月の予測値は2003年6月の実績値と予測値を使って計算しているだろう。実績値を使うのは当然としてもなぜ予測値を使うのかわかるか」

佐藤「前期の実績値を使うのはわかりますが、すでに終わって、しかもはずれた前期の予測値なんて意味がないと思います。競馬の予想でも過去のレースの実績は使いますが、過去のレースの予測なんて使いませんもん」

土屋「そこが競馬の予想と数学のちがいだ。ところで2003年6月の予測値はどうやって出している？」

佐藤「2003年5月の実績値と予測値を使っています」

土屋「ということは2003年6月の予測値には2003年5月の実績値も結果として使っているよなあ。つまり2003年7月の予測には前月だけでなく5月という前々月の実績値も加味されているわけだ。じゃあもう1つの2003年5月の予測値は？」

佐藤 「2003年4月の実績値と予測値を使っています。2003年4月の予測値には2003年3月の実績値が加味され…。そうか、わかった。2003年7月の予測値には過去の38期分の実績値が少しずつ入っているんだ。かしこいなあ。これを使えばいつでも翌月の予測は2つの数字だけで計算できるんだ」

土屋 「まあいってみれば指数平滑法というのは、直近を意識しながら過去のデータを全部使った平均なんだ」(図2-8)

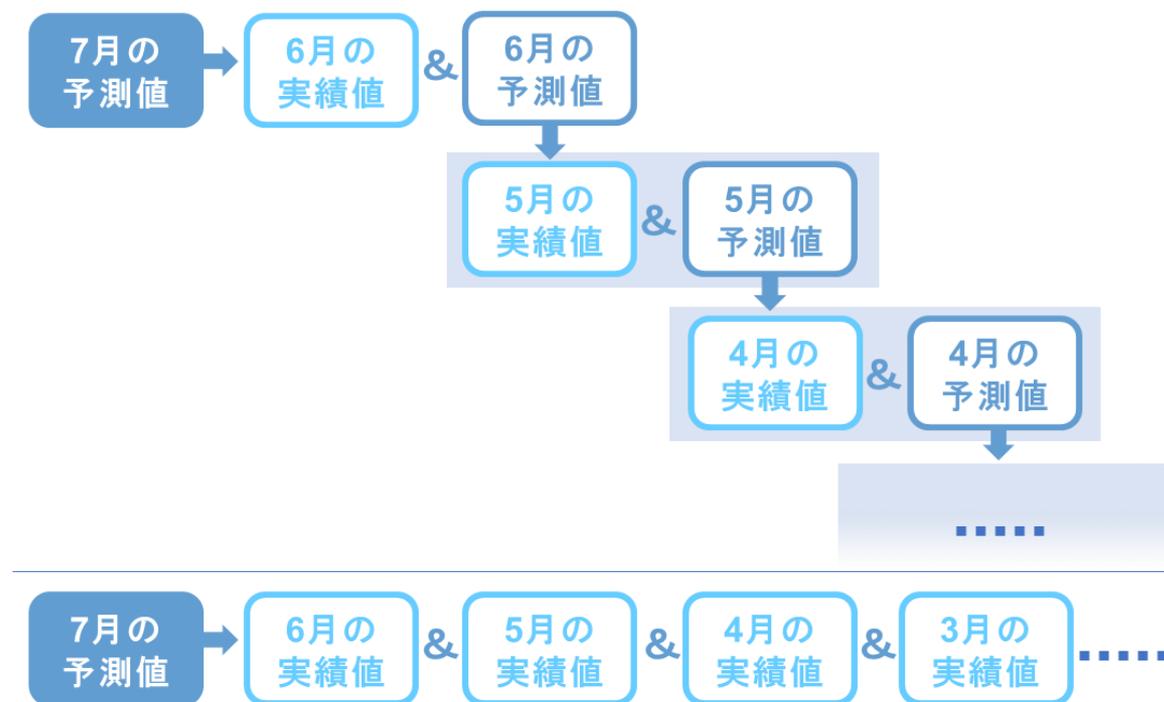


図2-8

佐藤 「ところで平滑化係数って何ですか。よくわからないので、本の例をまねして0.3を使ったのですが…」

土屋 「平滑化係数に0.3を使うということは、2003年7月の予測値を出す時、2003年6月の実績値に0.3、予測値に0.7をかけているわけだろう。6月の予測値は5月以前の実績値だ。つまり予測値の根拠が直前の月の実績が0.3、それ以前のすべての月の実績に0.7をかけているんだ。直近と過去のバランスが3対7ということだ」

佐藤 「その数字は、いくつにすれば良いのですか？」

土屋 「いくつということはないんだ。もちろん1を越えたらだめだけどな。これを大きくすれば直近月の実績値が大きく反映され、小さくすれば過去の値が大きく反映されるんだ」

佐藤 「どう決めたら良いんですか？」

土屋 「カンだ」

佐藤 「えーっ、カンを使うなら数学なんていらなんでしょう」

土屋 「そんなことないよ。需要予測というのはコンピュータが予測するのではなく、人間が予測するんだから、カンが大切だろう。もっといえばカンのない人は予測しちやいけないんだ。俺みたいにMCスーパーに行ったことがなく、スーパーグリーンがどう店頭に並んで、いくらで売られているかも知らない人間が予測したらだめなんだ。予測結果があたりそうか、あたりそうもないのかは人間が判断するんだ」

佐藤 「昔からヤマカンは得意でしたけど…」

土屋 「ヤマカンとは少しちがうんだ。今言ってるのはは根拠のあるカンだ。きつとこうなんだと確信するようなカンだ。これが前にも話した仮説だ。佐藤がセールスを続けていくなら、このカン、仮説を大切にしなきゃ。そのカンを活かす道具として、数学や統計学を勉強するんだ。すばらしいやり方を編み出した数学者も、カンというひらめきがあり、そのカンを数学で実証していったり、式で表していったんだ」

土屋 「ところで佐藤はこの予測値351個を見てどう思うんだ」

佐藤 「少し小さいと思います」

土屋 「佐藤はこの2003年7月はもっと売れると思うのか？」

佐藤 「だって5月で388個、6月で372個の実績で、7月を351個と予測するのは変ですよ。もう少し直近のデータを重視すべきだと思いますよ」

土屋 「だったら思い切って平滑化係数0.5でやってみろ。直近のウェイトを高めてみるんだ。こうすると7月の予測値の根拠は、6月が半分、5月以前が半分だ。ちょっと強烈だけどね」(図2-9)

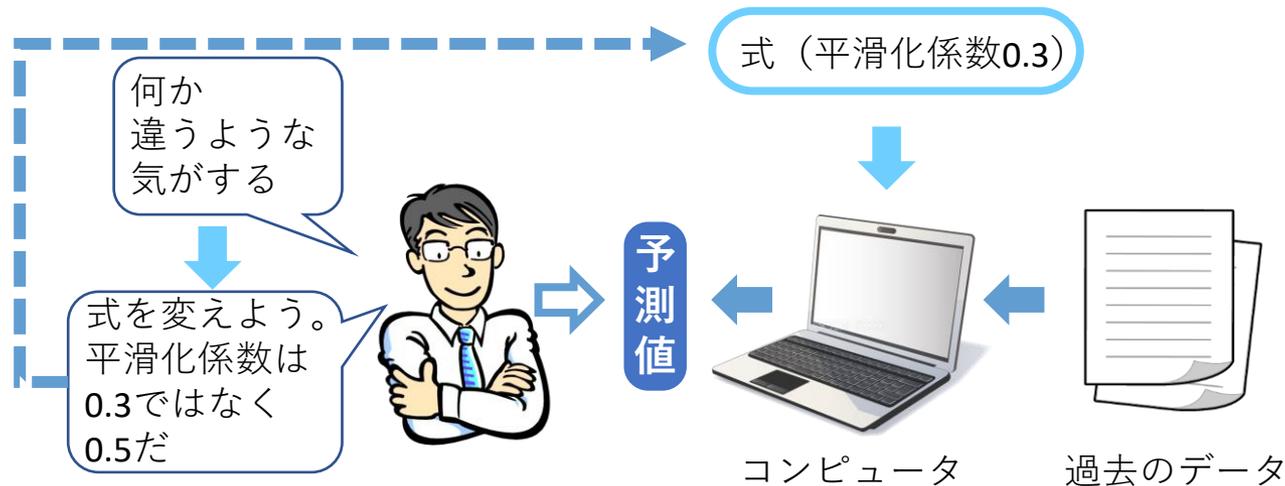


図2-9

- 佐藤** 「表計算ソフトで式を作っているので簡単です。ちょっと待ってください」
- 土屋** 「そうなんだ。表計算ソフトというのはもともとはこういうシミュレーションをやるために作られたものなんだ。それがいつの間にか単なる表や箱として使っている人が増えちゃったようだな」
- 佐藤** 「そういえばシミュレーションってよく使う言葉ですが、意味がはっきりわかりません」
- 土屋** 「シミュレーションとはコンピュータなどに実際の世界をまねして作り、それを人間がいろいろなことに使うことだ。大きなビルを建てるときに作る模型みたいなもんだ。ここでいえば表計算ソフトを使って、MCスーパーの店舗をコンピュータの中に作り、それを佐藤が予測に使うことだ。この時まず平滑化係数0.3ならどうなる、0.5ならどうなるとやって、人間が納得できる世界をコンピュータに作っていくんだ。これがシミュレーションのチューニングというやつだ。チューニングされたら、いよいよこのコンピュータの世界を使って予測していくんだ」(図2-10)

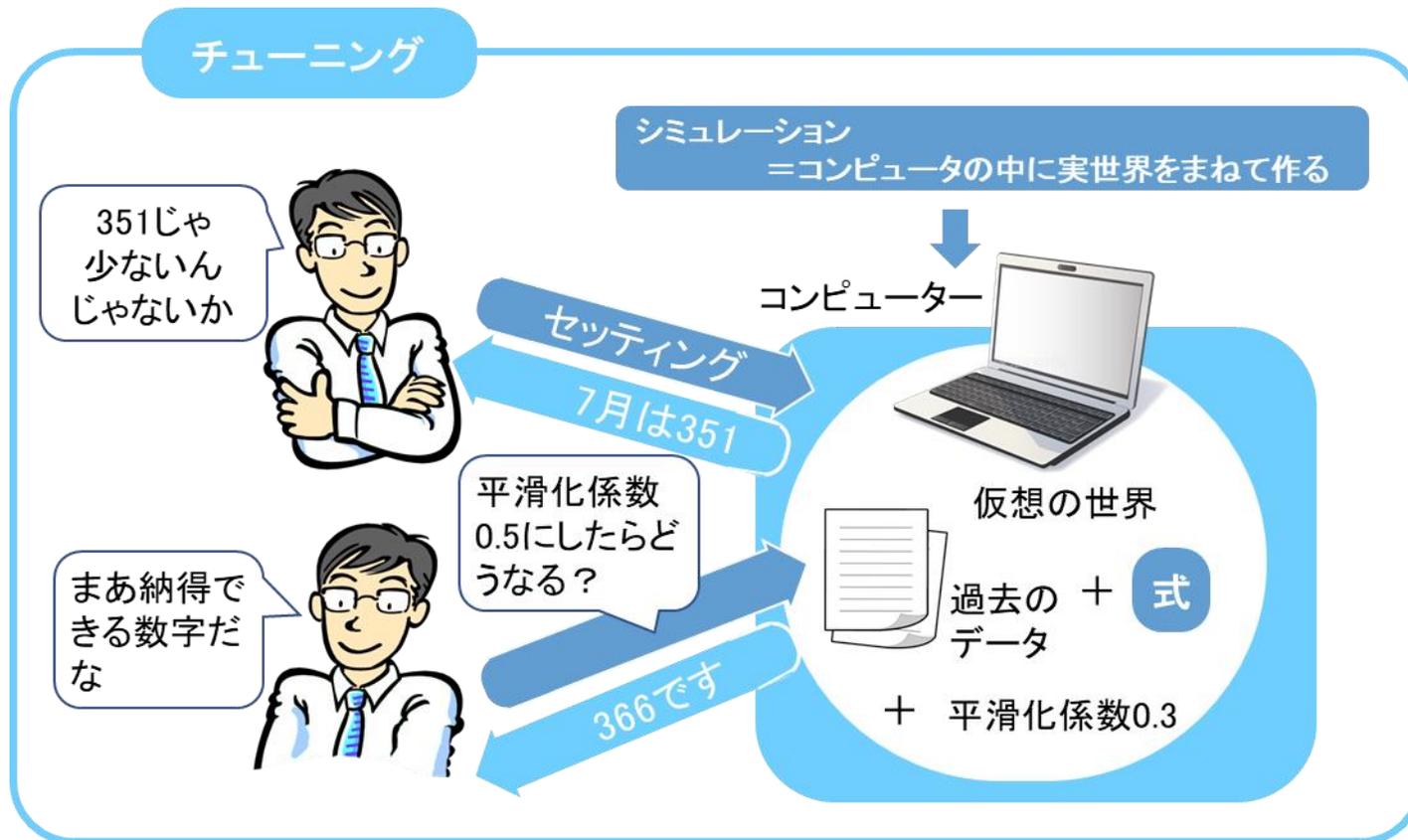


図2-10

土屋 「平滑化係数0.5で計算できたか」

佐藤 「出ました366です。まだ少ないようにも思いますけど、少し前の2003年1月が346で2月が298、3月が307あたりの動きを見ると、4月358、5月385、6月372で7月が366というのも納得できる線かもしれません」(図2-11)

平滑化係数=0.5

		実績値(D)	予測値(F)				
1期	2000年5月	120	120	21期	2002年1月	231	237
2期	2000年6月	126	120	22期	2002年2月	278	234
3期	2000年7月	142	123	23期	2002年3月	229	256
4期	2000年8月	130	133	24期	2002年4月	266	243
5期	2000年9月	128	131	25期	2002年5月	225	254
6期	2000年10月	146	130	26期	2002年6月	235	240
7期	2000年11月	205	138	27期	2002年7月	282	237
8期	2000年12月	182	171	28期	2002年8月	306	260
9期	2001年1月	224	177	29期	2002年9月	264	283
10期	2001年2月	206	200	30期	2002年10月	312	273
11期	2001年3月	188	203	31期	2002年11月	288	293
12期	2001年4月	201	196	32期	2002年12月	319	290
13期	2001年5月	198	198	33期	2003年1月	346	305
14期	2001年6月	246	198	34期	2003年2月	298	325
15期	2001年7月	233	222	35期	2003年3月	307	312
16期	2001年8月	211	228	36期	2003年4月	358	309
17期	2001年9月	240	219	37期	2003年5月	388	334
18期	2001年10月	222	230	38期	2003年6月	372	361
19期	2001年11月	213	226	39期	2003年7月		366
20期	2001年12月	255	219				

図2-11

土屋 「平滑化係数を0.5にするということは直近のデータをより大きく反映させるということだろう。どちらが良いか、どちらを使うか、平滑化係数をいくつにするかは人間が決めることなんだから、スーパーグリーンの売れ行きを一番良く知っている人、つまり佐藤が決めるべきだ。逆にスーパーグリーンのことを何も知らないのがコンピュータだ。自分で決めればその予測した数字をお客様に説明することも、なぜ平滑化係数を0.5にしたのかも説明できるだろう。そしてMCスーパーに意見を聞いて、予測値366で納得いかなければ平滑化係数を変えたり、他のやり方を使えばいい」

佐藤 「何となくわかってきました。予測式というのは人間の考えていることを数式を使って表現していくものなのですね」

土屋「そうだ。佐藤は0.3でも0.5でもいいから、数式を決めたら、なぜそうしたかを残しておくんだ。ここでは『より強く直近のデータを予測に反映すべきと思って、0.3から0.5にした』ということだ。これが残っていれば佐藤の後輩が予測値を見て、どうやって予測したかだけでなく、なぜそのようにしたかがわかるようになるんだ。これが今さわいでいるナレッジマネジメントだよ。佐藤が試行錯誤して得た、知識・ノウハウ・経験というナレッジを、誰でもわかる形に残しておく。そうすれば次の人がきっともっと良くしてくれるはずだ。佐藤だって先輩のナレッジがあれば良かったと思うだろ？」

佐藤「ここだけの話ですが、先輩がやった過去の予測値なんかを見たのですが、何をやろうとしているのか、どうしてこうなったかわかりませんでした」

佐藤はこの結果を中村支店長に報告しに行った。

佐藤「支店長、出来ました。指数平滑法を使って予測しました」

中村「おっ、勉強したな。366個か、ちょっと小さいような気もするなあ。ところで2003年8月以降はどうなるんだ。MCスーパーとは最低半年単位に、できればもっと長期に考えていくんだぞ」

佐藤「現時点では8月以降も366個と予測するしかありません。7月の結果が出たら変わると思います」

それを聞いて中村支店長の顔色が変わった。

中村「ちょっと待て。指数平滑法って何だ。説明してみろ」

佐藤は土屋に教わったことを説明した。

中村「それじゃスーパーグリーンではあたらないだろう。言ってみれば過去の平均値のようなものだろう。試しに実績値と予測値の誤差を出してみろ」

(図2-12)

平滑化係数=0.5

		実績値(D)	予測値(F)	予測誤差 (F-D)
1期	2000年5月	120	120	0
2期	2000年6月	126	120	-6
3期	2000年7月	142	123	-19
4期	2000年8月	130	133	3
5期	2000年9月	128	131	3
6期	2000年10月	146	130	-16
7期	2000年11月	205	138	-67
8期	2000年12月	182	171	-11
9期	2001年1月	224	177	-47
10期	2001年2月	206	200	-6
11期	2001年3月	188	203	15
12期	2001年4月	201	196	-5
13期	2001年5月	198	198	0
14期	2001年6月	246	198	-48
15期	2001年7月	233	222	-11
16期	2001年8月	211	228	17
17期	2001年9月	240	219	-21
18期	2001年10月	222	230	8
19期	2001年11月	213	226	13
20期	2001年12月	255	219	-36

21期	2002年1月	231	237	6
22期	2002年2月	278	234	-44
23期	2002年3月	229	256	27
24期	2002年4月	266	243	-23
25期	2002年5月	225	254	29
26期	2002年6月	235	240	5
27期	2002年7月	282	237	-45
28期	2002年8月	306	260	-46
29期	2002年9月	264	283	19
30期	2002年10月	312	273	-39
31期	2002年11月	288	293	5
32期	2002年12月	319	290	-29
33期	2003年1月	346	305	-41
34期	2003年2月	298	325	27
35期	2003年3月	307	312	5
36期	2003年4月	358	309	-49
37期	2003年5月	388	334	-54
38期	2003年6月	372	361	-11
	平均	240	227	-13

図2-12

中村「実績値と予測値で誤差の平均がマイナス13だろう。ということは毎月13個ずつ少なめに予測しているということだろう。今回の予測値366個だって、きっと13個少ないんじゃないか。佐藤はセールスで、評論家じゃないんだ。しっかりしろ。スーパーグリーンは成長期であり、過去の平均値より増加すると考えるべきだろう。去年マーケティングセミナーで勉強した商品ライフサイクル*を思い出せ」



商品ライフサイクル

商品が市場に登場してから消えてなくなるまでの時間的な推移を言う。多くの場合、時間を横軸、市場での総売上高を縦軸とするグラフで表す。

佐藤はマーケティングセミナーのテキストを取り出して、商品ライフサイクルについて調べてみた。

「商品のライフサイクルは4つのステージに分けられるか。そういえば人間の幼年期、青年期、熟年期、老年期に例えて講師が説明していたなあ」(図2-13)

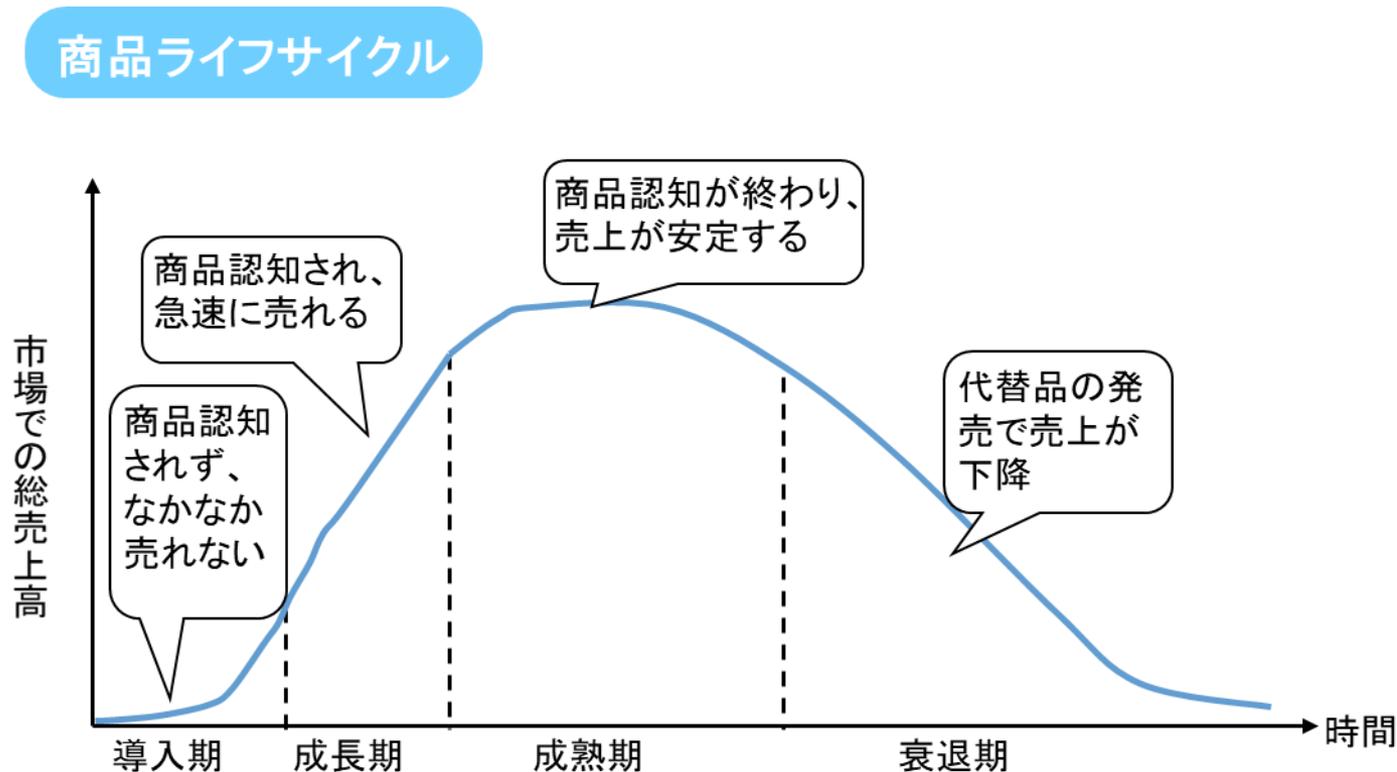


図2-13

佐藤はスーパーグリーンの過去の販売資料を取り出した。

「スーパーグリーンの発売は1998年7月か。私が入社する前だ。結構昔からあるんだなあ。2000年5月にマイナーチェンジ*してるんだ。あれ、この月にMCスーパーの中央店では店頭化している。先輩ががんばったんだなあ。資料にはスーパーグリーンはこのマイナーチェンジを機に、市場で伸び始めたと書いてあるぞ。ということは2000年5月から成長期か。まだ伸びているから、今でも成長期か。後ろに何か書いてあるぞ。『一般に健康飲料は成長期、成熟期とも長期間にわたるのが特徴であり、いわゆる息の長い商品である』待てよ、指数平滑法って要するに過去の平均をとってるんだよな。母集団は何だ。MCスーパー中央店のスーパーグリーンの月販個数だ。標本は過去の月販個数か。伸びているということは過去の月販個数と、これからの月販個数は明らかにちがう特徴を持っているんだから、別の母集団じゃないか。ということは指数平滑法の予測値、つまり過去の平均値という統計量を使うというのはあまり良い方法じゃないかもしれないぞ」



マイナーチェンジ

商品の機能・品質など的一部分の手直しを行うことを言う。大幅な変更はメジャーチェンジ、モデルチェンジと言う。もともとは自動車業界の用語。

佐藤「土屋さん、支店長に2つのことを指摘されてしまいました。1つは予測値が小さすぎることです。平均すると13も予測値の方が実績値より小さくなっています。もう1つは8月以降の予測値が出せないことです」

土屋「6月までの実績値が346、298、307、358、388、372だろう。まあ366も良い線だよな。おかしいなあ。その予測誤差が-13というのは本当か。ちょっと見せてみる」

佐藤は支店長に見せた表を出した。

土屋「なんだ、スーパーグリーンは着実に伸びているんじゃないか。うちにもそんな商品があったんだ。指数平滑法というのは、時間とともに変化していく数字の、「たまたま」を取ってなだらかにしていくという方法なんだ。これが「平滑」つまり「なだらか」の意味だ。基本的には過去と同じような動きをしめているものに向いているんだ。そう考えれば実は平滑化係数0.5というのは直前を重視しすぎて『なだらかさ』が弱いといえるんだ。2003年5月までの37個のデータと2003年6月のたった1つのデータを同じ重みで考えるというのは少しやりすぎだろう。本来なら0.2か0.3くらいが妥当なところだなあ。」

土屋 「もう1つの支店長の指摘の8月以降については、指数平滑法というのは昨日と同じ傾向と考えるんだから、7月の予測値と同じと考えるしかないんだ」

佐藤 「ということは今回の場合、指数平滑法はやめた方が良いということですか？」

土屋 「君はどう思う？」

佐藤 「やめたいです」

土屋 「じゃ、ちがう方法にしよう。予測するのは佐藤だ」

佐藤 「ちがう方法にはどんなものがあるのですか」

土屋 「まずこの数字からどんなイメージがわいてくる？」

佐藤 「私はスーパーグリーンは商品ライフサイクル上の成長期にあると思うんです」

と、佐藤はマーケティングテキストの商品ライフサイクルのカーブを見せた。

土屋 「へー、こんなものがあるんだ。品質管理ばかりやってきたんで、知らなかったよ。若い時マーケティングの勉強しなかったからなあ。いわれてみれば商品って生まれてから死ぬまで、こんなカーブになるような感じがするよ。佐藤はこの成長期の線が欲しいんだろう」

佐藤 「はい」

土屋 「だったらまずこのライフサイクルのカーブと同じように、さっきの38個のデータを時間を横軸に、販売個数を縦軸にしてプロットしてみろ」

と、紙を取り出した。

佐藤 「プロットって何ですか」

土屋 「1期目が120個なら、横軸が1、縦軸が120の所に点を打つということだ」

佐藤は定規を使って紙に目盛りをとり、点を打っていった。(図2-14)

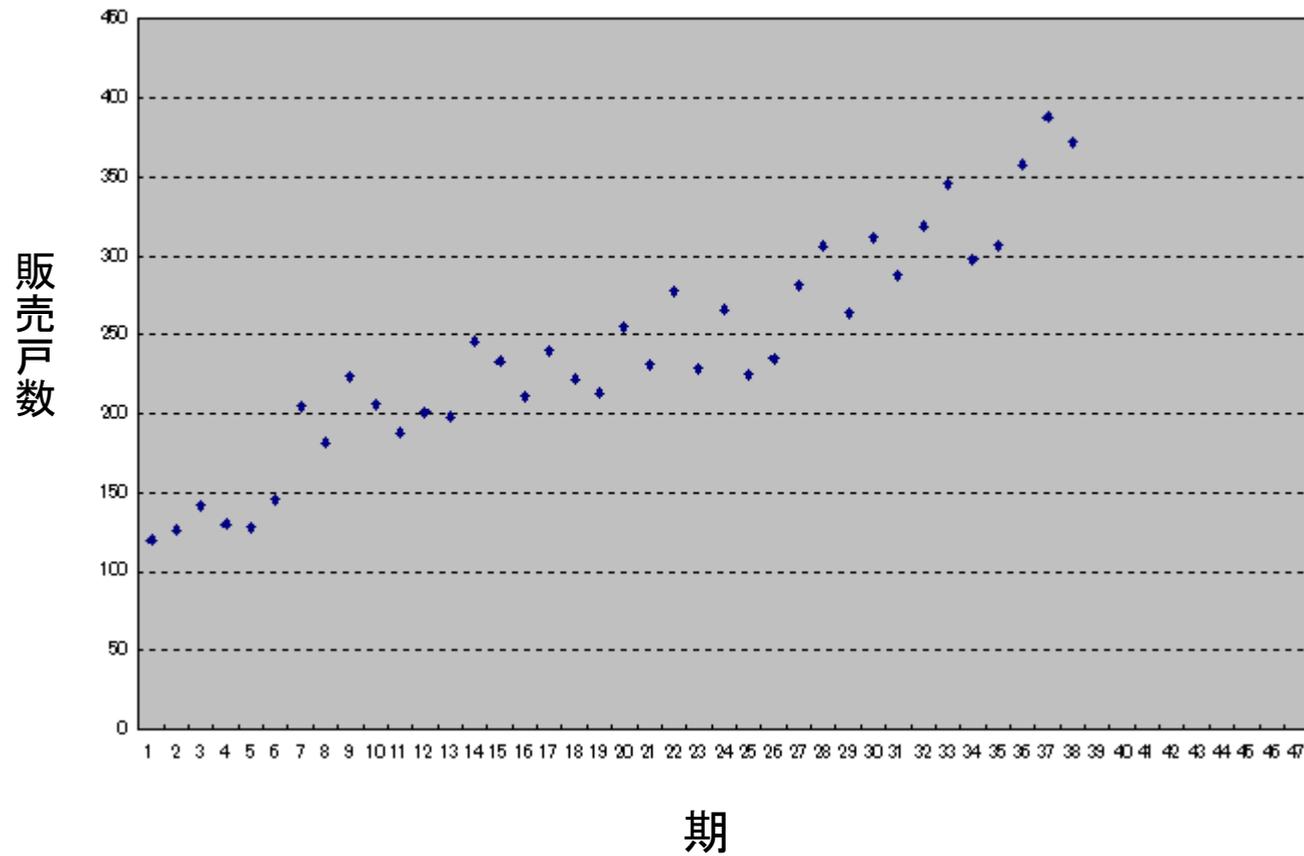


図2-14

土屋 「何か見える？」

佐藤 「どう見ても右肩上がりですよね」

土屋 「右肩上がりの直線が見えるか？」

佐藤 「なんとなく見えます」

土屋 「カンがいいな。じゃ、その線をアバウトに引いてみろ」

佐藤 「こんな感じかなあ」

佐藤は右肩上がりの直線をさっと定規で引いた ([図2-15](#))

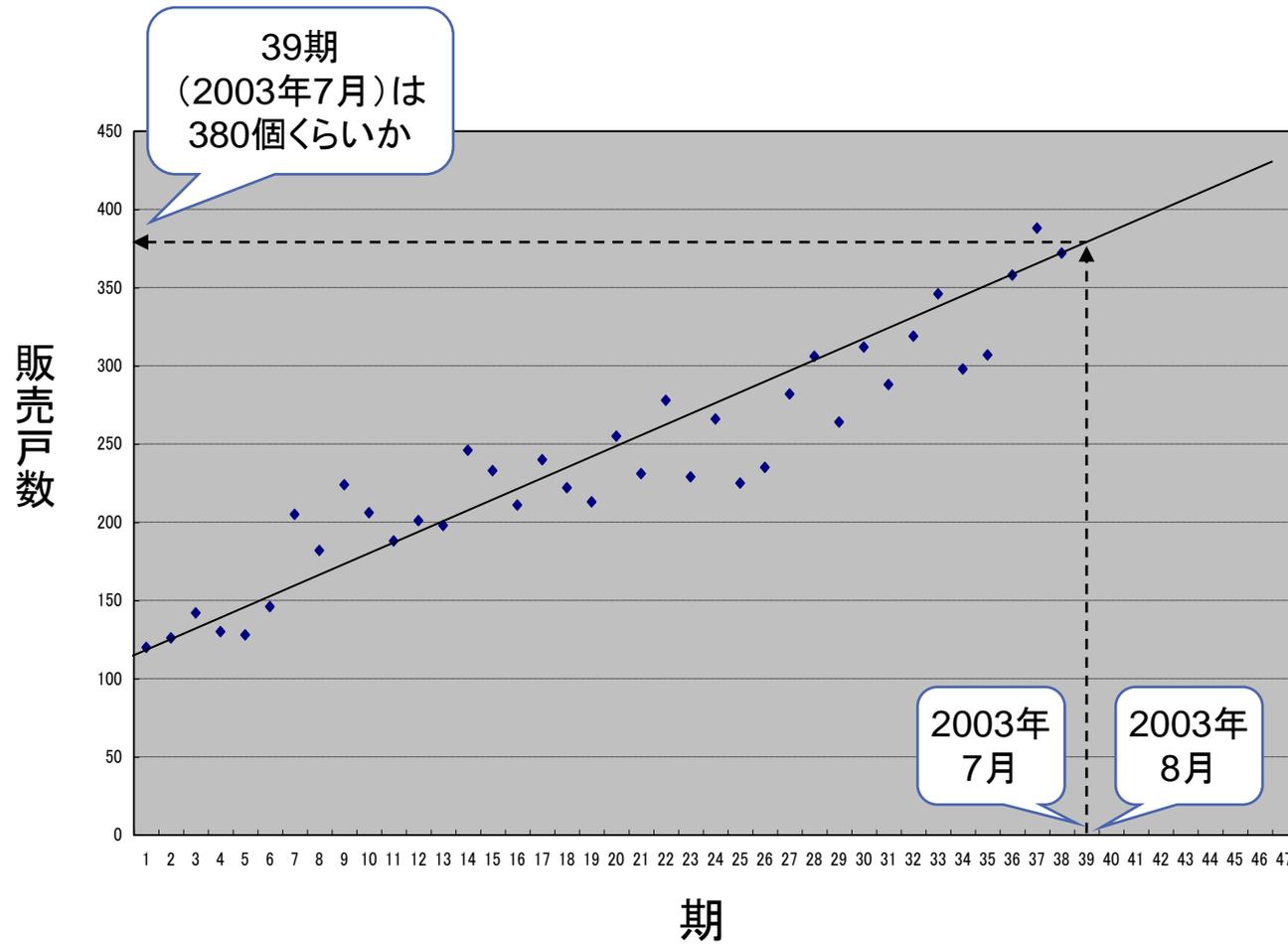


図2-15

佐藤 「そうかこの線で39期、つまり2003年7月を予測できるし、8月も9月もそれから先もずっと予測できるんだ。ちなみにこうやってみると2003年7月は380個くらいか。でも380個くらいじゃしょうがないですよ。できればコンピュータ使って、この『くらい』を取り、はっきり何個って出したいです」

土屋 「今、その線どうやって引いた？」

佐藤 「点と点の間ぐらいをねらって、さっと引きました。」

土屋 「それじゃお客様に根拠を説明できないだろう。しかも引き方によって数字が変わるので、引き方でもめてしまうだろ。何か佐藤の線は傾きが大きいような気がするぞ」

佐藤 「傾きが大きいって伸びが大きいってことですよね。」

土屋 「まあ、営業マンなんだから、伸びを大きくしたい気持ちもわかるけどな」

佐藤 「MCスーパーの木村さんにも同じこと言われそうで、やだなあ。そうすると木村さんが自分で線を引くだろうなあ。そして傾きでもめるだろうなあ。何か良い方法ないですか」

土屋 「それが数学なんだ。とびっきり頭が論理的な人が、他の人にぐうの音も出ないように理論化するのが数学さ。だから俺は好きなんだ。この時どうするのがもっとも論理的かと考えれば、佐藤のプロットした38個の点から直線までの距離の和がもっとも小さい直線を引けば良いんじゃないか。これをコンピュータにやらせればいいんじゃない」(図2-16)

佐藤 「そういえば自分も同じことをして線を引いたような気がします」

土屋 「コンピュータを使うというのは、ただ計算を早くやるんじゃない、実はここにも意味があるんだ。人間はファジーにさっと目分量で直線を引けるけど、コンピュータはそれが出来ないんだ。逆にいえば人間はさっと引いてしまうから答えの直線が1通りにならず、線の引き方でもめるんだ。コンピュータは頭が悪いから、きちんとやり方を1通りに決めて教えてやらないとできない。だから結果はいつも1通りだ。そうすれば引き方でもめない。コンピュータ化には、仕事の標準化を必要とするので、コンピュータ化すれば標準化の効果も上がるんだ」

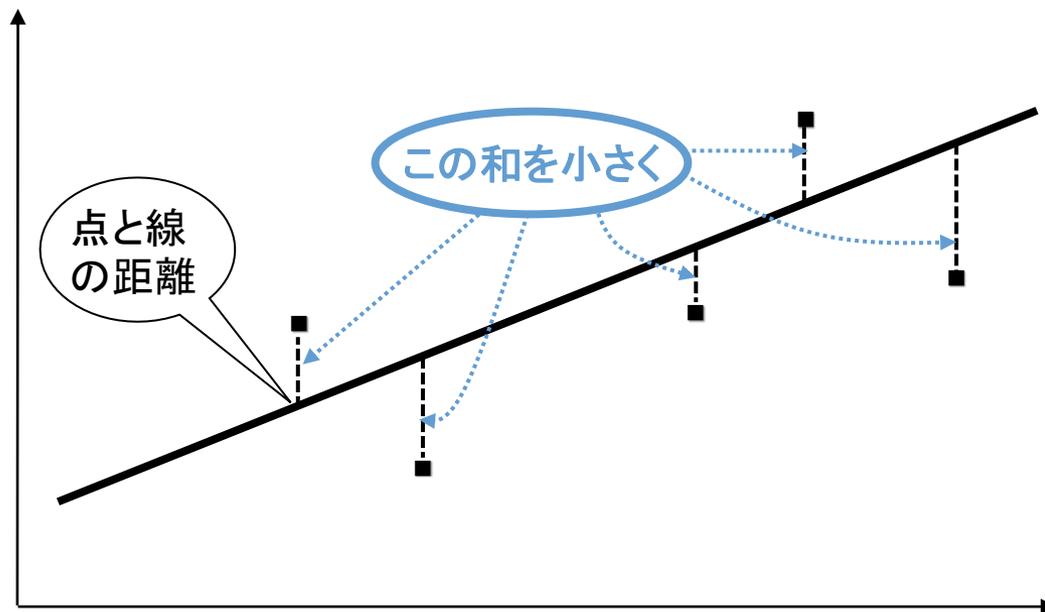


図2-16

土屋 「『距離の和を最小にする』方法を、人間が教えて、コンピュータが計算するんだが、ただ単純に距離を計算すると上にある時、下にある時でプラスとマイナスになってややこしいんだ。そこで距離をすべて正にするために、全部2乗してから足すんだ。数学の常套手段だ。覚えているか、標準偏差でやったろう。そのうえでこの距離の2乗の和の最小値を出せば良いわけだ。だからこういうやり方を最小2乗法っていうんだ」

佐藤 「ようするに距離の和を最小にするということですね。ところでどうやって最小値を出すんですか？」

土屋 「距離の2乗の和を微分するんだ」

佐藤 「えっ、微分ですか…。まいった…」

土屋 「微分なんてコンピュータがやってくれるから心配するな。自動車がどうやって動くか知らなくても運転できるだろう。表計算ソフトで簡単にやってくれるから。でも『なぜ最小値を出すのか』という理由は知っておかないと意味がないし、顧客に説明できないからね」

佐藤 「コンピュータがやってくれるということは、7月、8月、9月…と全部予測してくれるんですか」

土屋「予測してるのは佐藤で、コンピュータは佐藤の指示に従って計算するだけだ。佐藤が点と直線の距離の和が最小になるようにコンピュータに線を引かせ、予測値を出しているんだ。表計算ソフトを使えば、計算はあっという間にできるよ」

佐藤「やってみます」

佐藤は先ほど手でやった「プロットする」「直線を引く」「7月～12月の数を読む」ということを表計算ソフトを使ってやった。(図2-17・図2-18)

期	年月	予測値
39期	2003年7月	355
40期	2003年8月	361
41期	2003年9月	367
42期	2003年10月	373
43期	2003年11月	378
44期	2003年12月	384

図2-18

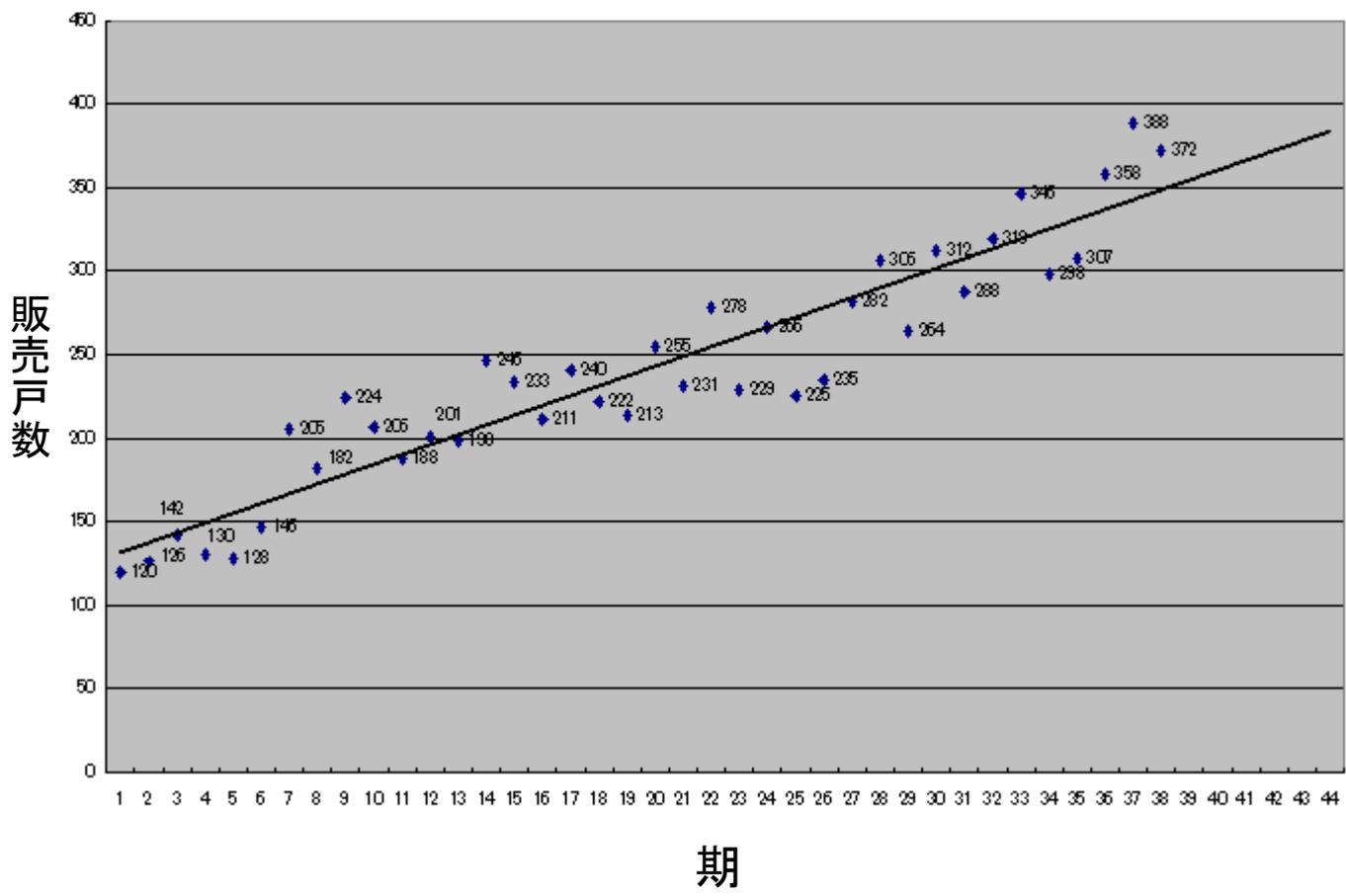


図2-17

佐藤 「7月の予測値は355か。あれ、さっきより減っちゃった。でも8月、9月と伸びて行って、いい感じかもしれないな」

土屋 「いい感じってどういう意味だ」

佐藤 「お客様に説明しやすい感じですよ。でもこのやり方でお客様が納得してくれますかね？あたららないと言いませんか？」

土屋 「そりゃ、あたららないかもしれないよ。ライバルだっけていて、似たようなもの出すかもしれないし、世の中だっけて変わって、急に健康ブームが去っていくかもしれないよ」

佐藤 「それじゃ困るな」

土屋 「あたららないと困るっていうなら、予測なんてやるな。占い師に聞けば」

佐藤 「ずいぶん乱暴な。だっけて『あたららない予測』なんて意味ないでしょう」

土屋 「そうかあ？あたるかあたららないかは、やってみないとわからないだろう」

佐藤 「何だか禅問答のようですね。昔から理系の人と話すといつもこうなってしまうんですよ」

土屋「整理してあげよう。予測というのは常に過去のデータからやるよな。この過去のデータをもとに何らかの加工をして、予測値を出し、これを人間が使うわけだ。つまり予測というのは過去のデータ、予測のやり方、予測値そして人間から成っているわけだ。予測値を見て人間が納得いかない数字が出たら、予測のやり方を変えるんだ。そしてなぜそれを変えたかを残しておくわけだ。さっきシミュレーションの時、話したろう」(図2-19)

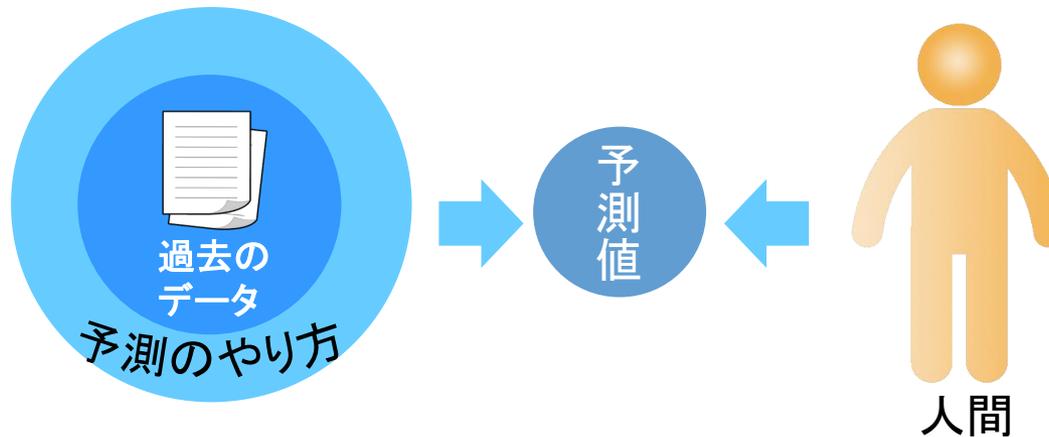


図2-19

佐藤「そうかあ。予測のやり方が指数平滑法であり、最小2乗法か」

土屋「人間がこの予測のやり方で予測値を納得できないときは答えは2つしかない。予測をしないか、予測のやり方を変えるかだ。まさか過去のデータを変えようとは思わないよな」

佐藤「指数平滑法で出した予測値に、人間である自分が納得いかないから、やり方を最小2乗法で直線を引くことに変えたんだ。同じようにお客様が予測値に納得しないなら、予測のやり方を変えれば良いんだ。自分が心配してたのは予測値の精度という得体の知れないものだったんだ。お客様と話すべきことは、予測値という結果だけではなく、やり方というプロセスの検証が必要なんだ。プロセスが決まれば結果は自ずと決まるわけだ」

土屋「前から思っていたんだが、どうしても文系の人間は予測値を見て、あたるかあたらないかを議論するけど、それって変だろう。予測のやり方を議論すべきだろう。納得いかないからといって自分の気分で予測値を変えるのは乱暴だろう。自分だけで責任取れるようなことなら良いけど、企業で何かを予測するときはもちろんのこと、お客様と予測値を共有する時は、予測値よりも予測のやり方を決めることが大切だとわかるだろう」

佐藤「わかります」

土屋「やっこの支店にも俺の言っていることをわかってくれる人が出てきたよ」

佐藤「会社に入ったばかりのときの社長の講話で、『これからのビジネスマンは先の読める人間になれ』といわれたことをよく覚えています。その時は予測力のようなものがある、仕事をしていくうちに自然に身についていくのかと思っていました。予測力とは予測のやり方をいくつ持っているかなんですね」

土屋「ただし誤解するな。予測のやり方は人間のカンがベースで、それを式で表していくことだ。そしてカンだけの時と違うのは、なぜそういう予測をしたかを説明できることだ」

佐藤は一週間後、MCスーパーの木村バイヤーと打合せをしていた。

佐藤「2003年7月から12月まで向こう6ヵ月間のスーパーグリーンの販売予測はお手元の資料にあるとおりと考えます。商品ライフサイクル上、スーパーグリーンは成長期にあり、右肩上がりです。このトレンドを最小2乗法によってお手元のような直線で表し、予測しました」

木村「こんなのあたるの？ミドリ食品の希望的観測じゃないの？」

佐藤「あたるかどうかはわかりません。ただ過去のデータから考えて、こうなるのが妥当という線を出してみました。木村さんが前におっしゃった『普通にやれば売れる数字』だと思っています。ただしスーパーグリーンが成長期にあって過去と同じような傾きを示していくことが前提です」

木村「ここまで伸びてきたんだから、明日は伸びないと考えるのも店長に説明しづらいな。当面はこれでいくか。でもいつか成熟期に入っていくんじゃないの」

佐藤「これからは毎月、実績値が出たら、それによって直線を引き直していきましょう。この直線の傾きが落ちてきて、つまり伸びが止まったら成熟期と判断し、予測のやり方を変えたいと思います」

木村「どうするの？」

佐藤「その時点からは指数平滑法を採用したいと思っています」

木村「それって何？」

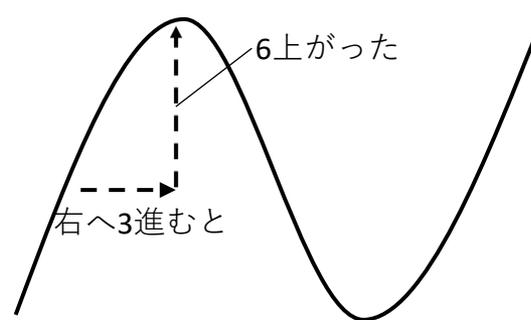
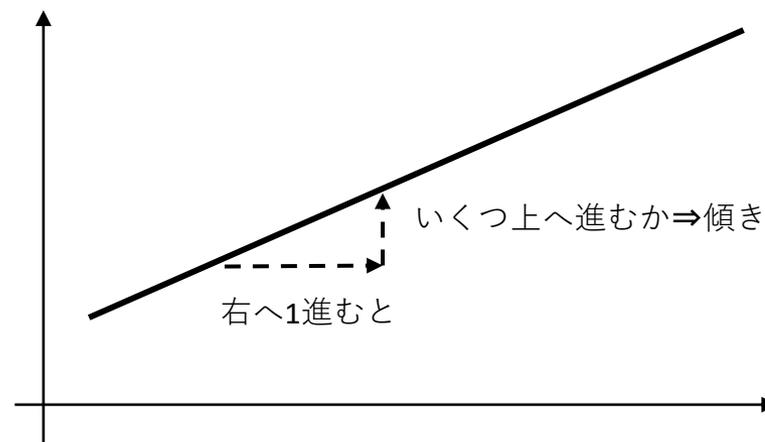
佐藤「過去のデータを直前に重さをおきながら、平均をとっていくものです」

微分

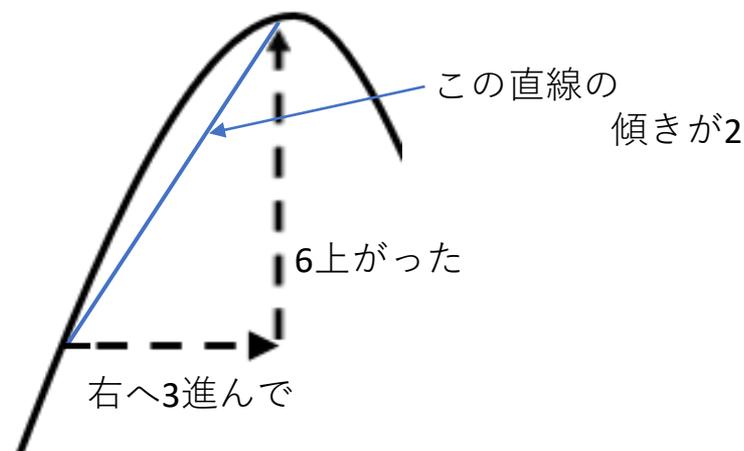
数学では「傾き」という表現をよく使います。これは横軸の数字が右へ1進むと、縦軸が上へいくつ進むかで表わしていきます。

右へ1進むと2上がるなら、傾きは2です。逆に右へ1進むと3下がるなら傾きは-3となります。この傾きがいつも一定のものを直線といいます。「傾き2」の直線では右へ3進むと上へ6上がり、右へ10進むと20上がります。

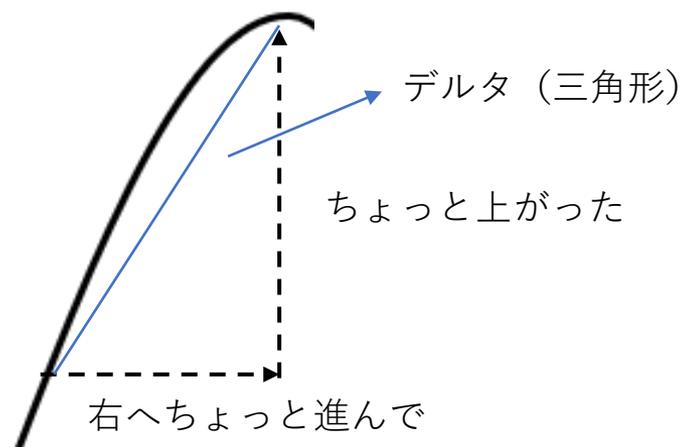
傾きを曲線で考えて見ましょう。曲線では場所によって傾きがちがいます。曲線上のある「点」に着目してみます。



この曲線では、図の点から、もし右へ3進んで6上がったら、その傾きは2です。その部分を拡大してみましょう。



この右へ進む幅をどんどん小さくして、その部分を拡大してみるとどんな風になるでしょうか。



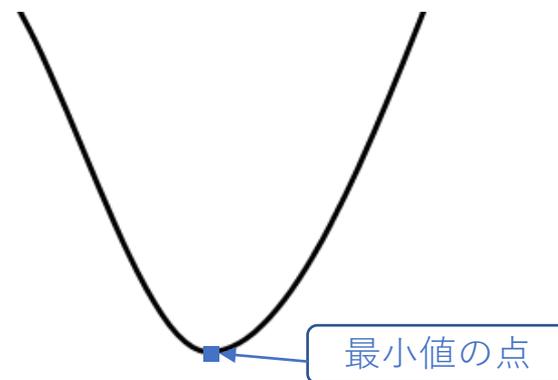
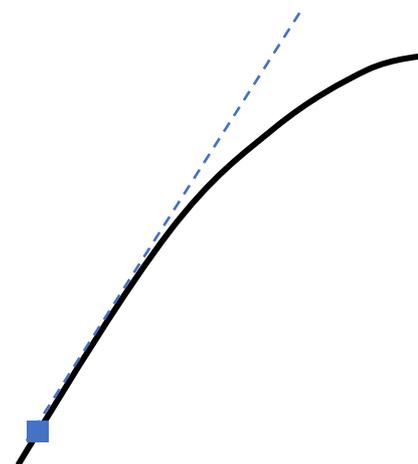
この「右へちょっと行く」を限りなく小さくしていくと、この点で曲線に接するような直線ができます。

これが接線です。この接線の傾きが、この曲線のこの点での微分係数といわれる数字です。いってみればこの点における傾きのようなものです。傾きはプラスならば増加、マイナスならば減少の傾向にあります。微分係数とはこの点において増加傾向か減少傾向かを表しています。

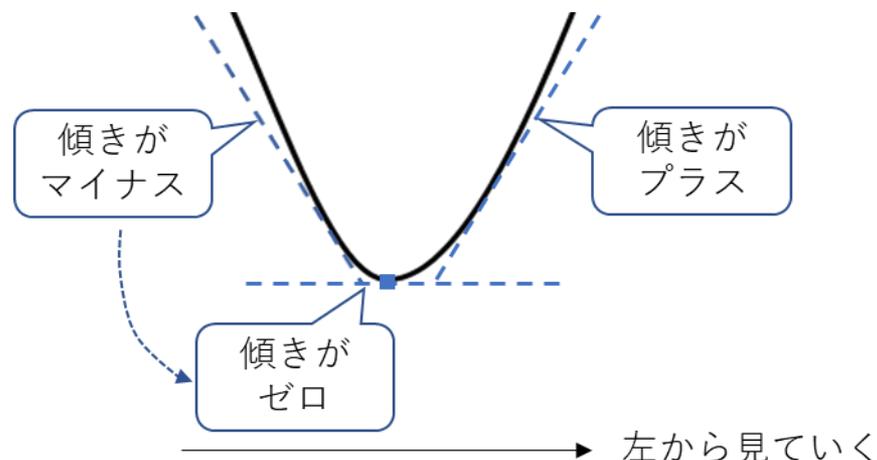
微分という言葉はこの「右へちょっと行く」を、「限りなく小さくする」という意味です。先ほどの図の小さな三角形のことをギリシャ語で「デルタ」といい、 Δ で表します。

よく微分に dx のように、 d を使うのは、*delta* の d のことです。この三角形の斜辺の傾きが微分係数です。

次に最小値について考えましょう。



最小値の点の周りの傾きはどうでしょうか。左から見ていって下さい。



最小値より左側では傾きがマイナス(減少傾向)ですが、右側ではプラス(増加傾向)なのがわかると思います。そして最小値の点では傾き、つまり微分係数が0(増加も減少もしていない、つりあった点)です。最大値の場合も考えてみて下さい。傾きがプラスから(増えていって)、0となり(最大値)、次に傾きがマイナス(減っていく)になっていくはず です。

つまり傾きが0、つまり微分係数が0の点では最小値か最大値をとります。本文で「距離の和の最小値をとるとき微分して求める」というのは、この微分係数が0の点を探すという意味です。

第 3 章

新しい店の売上を考える

MCスーパーはS県内のT市に新規店舗を出す準備を進めていた。このT店の店長には中央店の木村バイヤーが昇格することが内定しており、木村氏の後任のバイヤーには木村氏の下で働いていた松本氏が昇格することになっていた。木村氏はバイヤー業務の引継ぎとともに、本社で毎日のように開かれる新店準備委員会にも出席する忙しい毎日を送っていた。

木村氏は今頭を痛めていた。本部長からT店の年間売上高を予測しろと言われているのだが、どうして良いかわからずにいた。MCスーパーはもともとは青果店であり、スーパーマーケットへの業態転換*後30年たっていた。しかし近年は大店法*の影響で新規出店はなく、T店はMCスーパーとしては、実に10年ぶりの出店である。最後の出店を担当した店長も今や退職しており、新規店舗の売上予測方法など、誰に聞いていいかもわからずにいた。

そんな時、この1年間木村バイヤーとともに色々なことにチャレンジしてきたミドリ食品の佐藤が、中央店に来店した。

 業態転換

小売業においては業種が「何を売るか」をさすのに対して、業態とは「どう売るか」をさす。業種では青果を売っていれば青果店、鮮魚であれば鮮魚店と言う。業態とはスーパーマーケット、百貨店、コンビニエンスストア、一般小売店などの売り方をさす。この業態を変えることを業態転換と言う。

大店法

「大規模小売店舗における小売業の事業活動の調整に関する法律」のこと。スーパーなどの大型小売店の出店を規制し、中小小売業を守るための法律。大店法の下では好立地の商圈に大型店を出すのが難しかった。現在は大店法は廃止され、代わって環境規制のための大店立地法（大規模小売店舗立地法）が施行されている。

佐藤 「おめでとうございます。いよいよ店長ですね。T店も私の担当エリアですので、今後ともよろしく申し上げます。困ったことがあったら何でも言って下さい。すぐにかかけつけます」

木村 「もうすでに困っているんだ。助けてくれないか。実は新店の売上高を予測しなくてはならないんだけど、どうしていいかわからず困っているんだ。知ってのとおりうちの本部長は数字に細かいんで、いいかげんなことをいったら店長昇格どころかクビになっちゃうよ」

佐藤 「予測なら今まで何度も一緒にやってきたじゃないですか。予測するための過去のデータはありますよね」

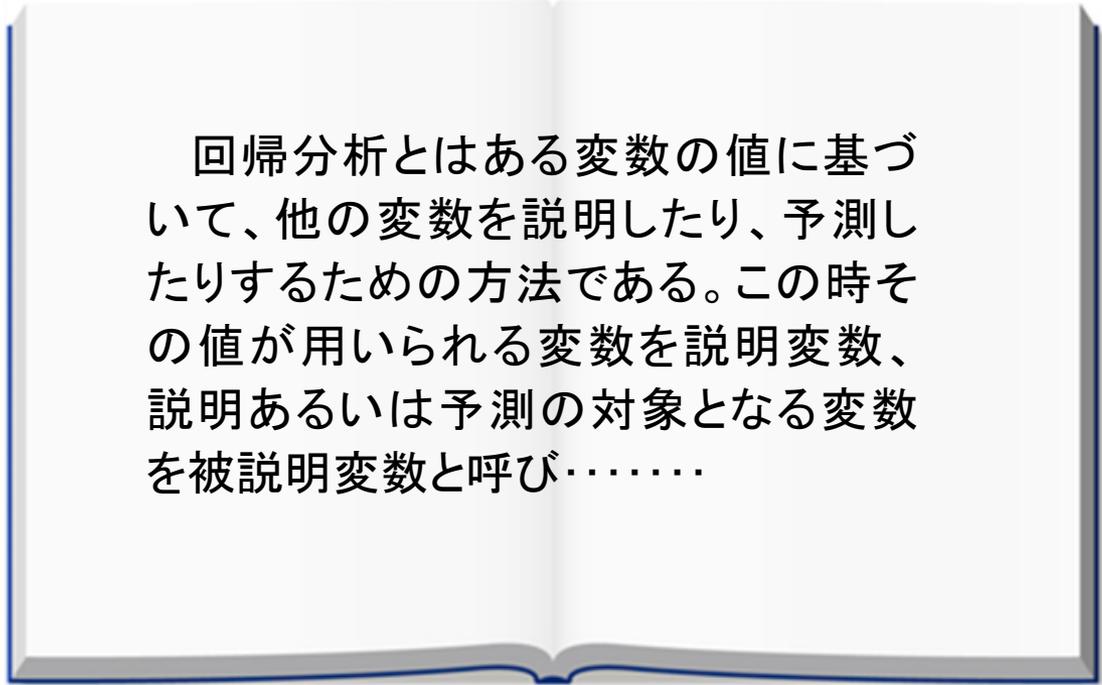
木村 「過去のデータといっても新規出店なんだから、今度の店のものはもちろんないよ。本部からもらったのは、うちの40店の昨年の年間売上高と各店の属性データなんだ。本部長はこれを使って予測しろっていうんだ」 (図3-1)

佐藤 「わかりました。調べてみます」

店名	前期売上高(百万円)	店舗面積(m ²)	商圏内人口	商圏内20~40代女性人口	商圏内世帯数	取扱商品アイテム数	生鮮(惣菜含む)アイテム数	加工食品(飲料含む)アイテム数	日用雑貨アイテム数
東川松店	604	1100	24,564	8,088	10,284	8,980	2,200	3,850	1,420
西川松店	426	720	18,226	6,425	8,726	7,420	2,050	2,670	1,520
⋮		⋮		⋮					
⋮		⋮		⋮					
T店(予定)	—	1,485	22,546	7,650	9,526	10,000	3,000	4,000	2,000

図3-1

「これって『誰でもわかる統計の本』の予測の所の、最初に書いてあったやつだ。確か時間以外の変数を使う回帰分析とかいうやつじゃないか」



回帰分析とはある変数の値に基づいて、他の変数を説明したり、予測したりするための方法である。この時その値が用いられる変数を説明変数、説明あるいは予測の対象となる変数を被説明変数と呼び……

図3-2

「説明変数と被説明変数のちがいがよくわからないぞ。そもそも何で回帰というんだ？回帰って『戻る』という意味だよな」

翌日佐藤は土屋に質問した。

佐藤 「回帰分析をやりたいんですが、よくわかりません。本にはビジネスパーソンのこづかいと年収のことが例として書いてあるのですが・・・」

土屋 「回帰分析か、もう佐藤は何度もやってるんだけどなあ。まあいいや。ところでこづかいと年収のうち、どちらが知りたいんだ」

佐藤 「ある企業がビジネスパーソンの年収を知りたいんだけど、年収を直接アンケートで聞くのが難しいので、アンケートでこづかいを聞いて、その結果から1人1人の年収を推定すると書いてありました」

土屋 「だったらこの場合、年収が被説明変数で、こづかいが説明変数だ。こづかいで年収を説明したり、予測したりするんだ。説明は計算といってもいい、こづかいで年収を計算するんだ。年収はこづかいで説明されるから被説明変数、こづかいは年収を説明するから説明変数だ。『被』とは『される』ということだ。裁判の原告と被告だ。訴えるほうが原告、訴えられるほうが被告だ」

佐藤 「こづかいを聞けば年収がわかるんですか」

土屋 「それだけでわかるわけじゃないか。まず年収とこづかいの両方がわかっている人がいるはずだ。そんな表なかったか？」

佐藤 「ありました。これですね」(図3-3)

土屋 「10人か。随分少ないけどまあいいや。母集団は日本全国のビジネスパーソンで、標本はこの表にある10人のビジネスパーソンだ。この10人の年収とこづかいという2種類の標本値を使って分析するんだ」

佐藤 「統計量は何ですか」

土屋 「これはちょっとちがうんだ。こづかいと年収の関係を式で表すんだ。いってみればこの式が統計量みたいなもんだ」

佐藤 「その式を何に使うんですか？」

土屋 「この10人のビジネスパーソンのデータからこづかいと年収の関係を式にしておけば、こづかいがわかれば年収が計算できるだろう」

NO.	こづかい	年収
1	4.5	520
2	5.0	600
3	10.0	950
4	3.0	420
5	8.0	750
6	6.0	700
7	6.5	620
8	5.0	630
9	3.0	400
10	6.0	650

図3-3

佐藤 「そうか、どんな日本のビジネスパーソンでもこづかいを聞けば年収が推測できるということか。年収とこづかいの関係を式で表すなんて、『リンゴが落ちるやつ』えーと、そうだ万有引力の法則みたいですね」

土屋 「それはちがう。万有引力の法則はニュートンが引力と距離・質量の関係式をきちんと証明したから、いつでもその式であたっているわけだ。回帰分析はそういうものでなく、2つのデータの関係を手元にあるデータから推定して、その関係を式で表すんだ。10人のこづかいと年収のデータから式を作るんだ」

佐藤 「ということは、あたらないこともあるんですか？」

土屋 「あたるかどうかなんてわからないさ。だって年収があたっているかどうかわかるということは、本当の年収がわかったということだろう。だったら推定しても意味ないよ。そもそもさっきの10人だってその式どおりになるわけないさ」

佐藤 「この10人さえもあたっていないんですか」

土屋 「この10人の年収とこづかいをぴったり1つの式で表すなんて不可能だ。だってこの10人の中にこづかいが同じでも年収がちがう人もいるだろう。見てみる。こづかい3万円で年収420万円と400万円の標本があるだろう。ということはこづかいがわかっても年収は決まらないはずじゃないか。今あるデータから考えるとこづかい3万円の人はこれ位の年収と考えたほうが良いということだ。しかも回帰分析ではデータの数が増えればその式は変わるんだ。つまり年収とこづかいが両方わかっている人が10人の時と1000人の時では式がちがってくるんだ」

佐藤 「そうか統計量と同じですね。平均値だってそうだ」

土屋 「そのとおりだ。回帰分析とは平均値を出すのによく似ているんだ。前にやった紅白の視聴率も調査対象が変わったり、増えたりすれば変わるのとも同じだ。こう考えてもいいぞ。例えばこづかい5万円の人は平均して年収600万円と考えられるということだ」

佐藤 「回帰分析の結果の使い方はわかりましたので、やり方を教えて下さい。私が読んだ本には複雑な式が書いてあって意味がわからなかったんです・・・」

土屋 「式を考える前にグラフでやってみよう。その本にあるこづかいと年収がわかっているビジネスパーソンのデータが10個あるだろう、それをこづかいを横軸、年収を縦軸にプロットしてみろ」

佐藤は紙に定規を使ってプロットした。

佐藤 「前にもやりましたよね。できました」
(図3-4)

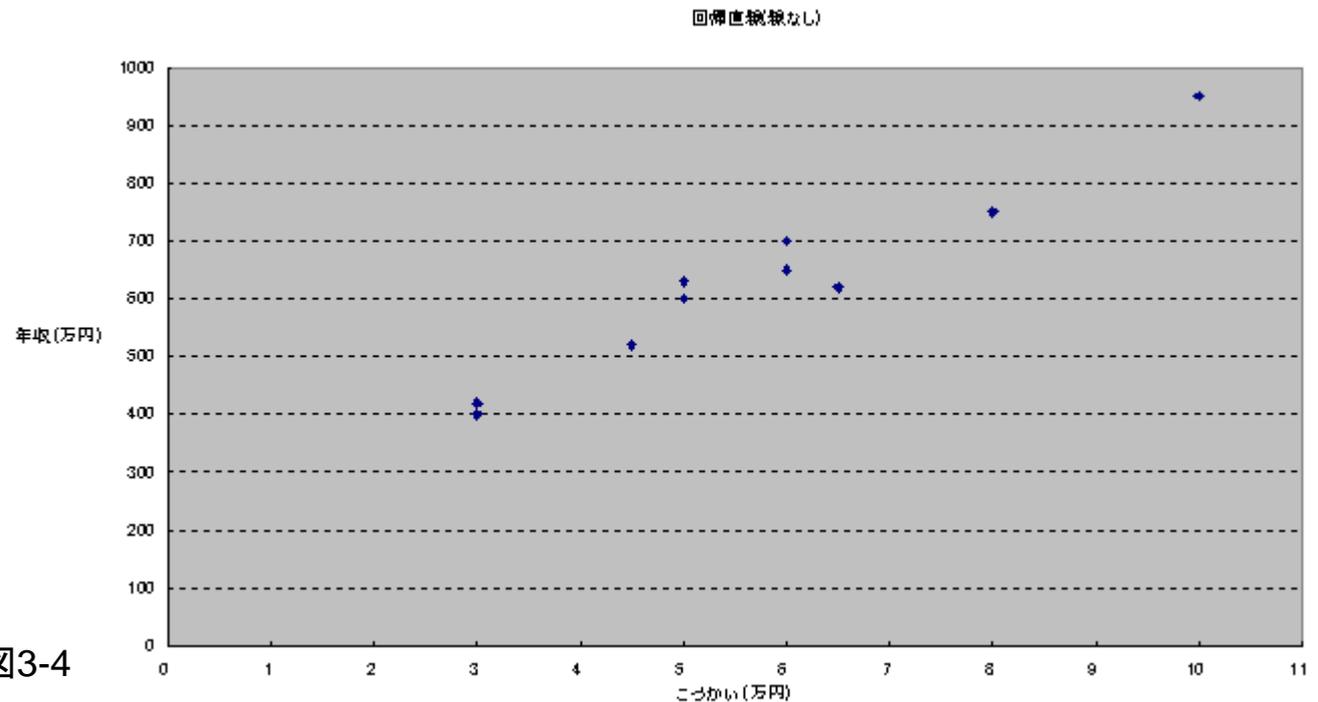


図3-4

土屋 「また何か線みたいなものが見えるか」

佐藤 「見えてきました」

土屋 「そのカンを大切にしろ。線が見えないときは回帰分析はやめた方が良い」

佐藤 「線を引きましょうか。各点からの距離の和、ちがった、距離の2乗の和が最小になるようにするんですよ。最小2乗法ですね」

土屋 「表計算ソフトを使ったほうがいいよ。そうすればこづかいがわかれば年収が自動的に計算できるようになる」

佐藤は表計算ソフトにデータを入れ、回帰分析の機能を使ってみた。あっという間にプロットされた直線が表示された (図3-5)

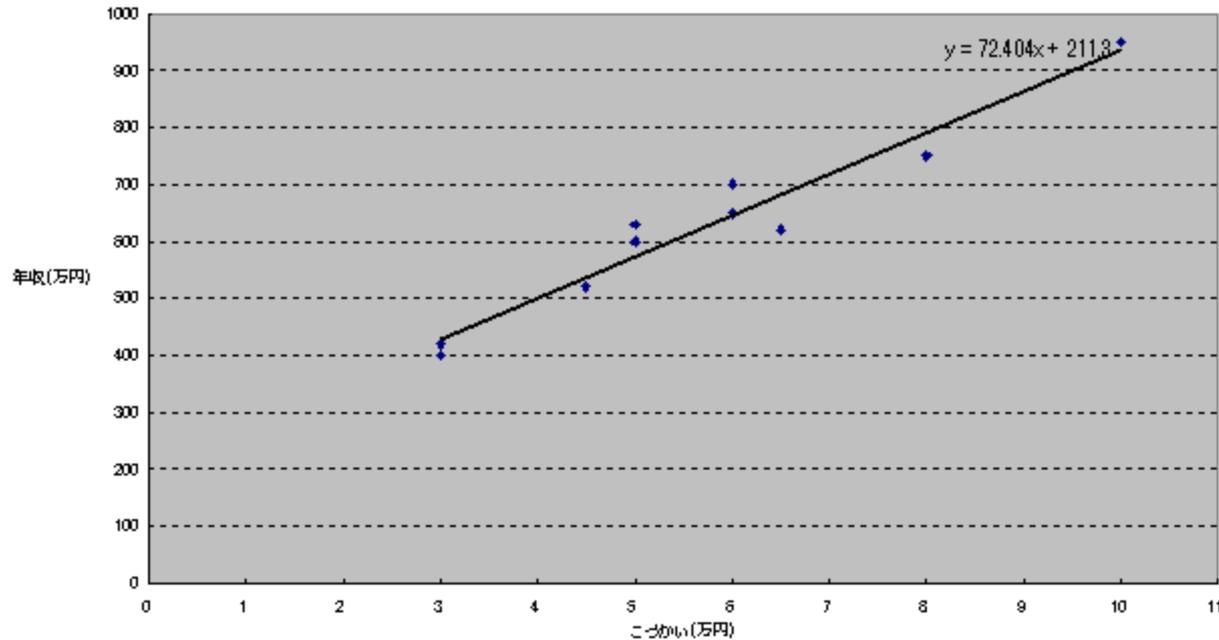


図3-5

佐藤 「この右上に出て来た式がさっきいった式のことですか？」

土屋 「そうだ。こづかいが x で年収が y だ。小数点以下は意味ないから無視していい。つまりこづかいを72倍して211足せば、年収が予測できる。こづかい7万の人は年収はいくらだ？」

佐藤 「715万円です。こづかい7万円の人の平均的な年収は715万ということですね」

土屋 「この直線が回帰直線というやつだ。そしてこれが回帰分析だ」

佐藤 「ということは前にやったスーパーグリーンの販売予測も同じように線を引いてるんですから、回帰分析ではないですか？」

土屋 「そうだ。あれはスーパーグリーンの販売額が被説明変数、時間が説明変数の回帰分析だ」

佐藤「ところで何で回帰というのですか？」

土屋「これは結構有名な話なんだ。イギリスのゴルトンという学者が親の身長と子供の身長の関係を調べていて、佐藤のようにこれをプロットしていったんだ。あたりまえといえばあたりまえだが、背の高い親からは背の高い子供が出来ていたんだけど、親の背の高さの差ほど身長は遺伝されていなかったんだ。ゴルトンはこれを人間の身長は世代が経つにつれて、人間の平均値へ戻っていくと考えたんだ。これが戻る、つまり回帰という意味だ。まあ言ってることはあまり正しくなかったんだけどね。たださっきの年収とこづかいのグラフを左から見ていってみろよ。直線の下に点があると、次は上にあったりして、各点がこの直線に戻っていくように見えるだろう。そんな感じがするから、今でも回帰という言葉を使っているんだ」

佐藤「なるほど。じゃあ、T店の予測にとりかかります。まず年収にあたる被説明変数はどう考えても年間売上高だな。ビジネスパーソン10人がMCスーパーの既存店だ。あれ、こづかいにあたる説明変数はたくさんあるぞ、どれにしたらいいんだ」

土屋 「直感で佐藤が決めてもいいけど、お客様が納得しないよなあ。こっちが良い、あっちが良いで、説明変数でもめたら予測の意味がないよな」

佐藤 「どうでしょう」

土屋 「そんなこと佐藤以外の人間も悩んでいるんだから、その中のもっとも賢いやつが答えを出しているはずだ。いつもそう考えるんだ。その答えが相関分析だ」

佐藤は帰宅し、早速「誰でもわかる統計の本」で相関分析の所を読んだ。回帰分析と同じ章で説明してあった。(図3-6)

相関分析とは2種類の母集団から標本値をとり、その関係を分析することで、当該2種類の母集団の関係を推定するものをいう。ピアソンの積率相関係数(r)がもっとも有名である。

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

図3-6

「回帰分析と同じでこづかいと年収の例を使っているぞ。こづかいと年収が2種類の母集団ということだな。そうか、ビジネスパーソンでなくこづかい、年収を母集団と考えるのか。そしてその関係を分析するのが相関分析か。あれ回帰分析だってこづかいと年収という2つの母集団の関係を分析して式にしてるじゃないか。回帰分析と相関分析はどこがちがうんだ」

翌日佐藤は土屋に質問した。

佐藤「相関分析と回帰分析ってどこがちがうんですか？」

土屋「文系にしては良いことをいうじゃないか。世の中はみなファジーに使ってるが、実は数学者で相関分析という表現を使う人はあまりいないんだ。あえて言うのなら、相関分析が『2種類の母集団の関係を推定する』と定義されるなら、回帰分析だって2種類の母集団の関係を推定しているんだから、相関分析の一種だ。本に書いてあったり、よく使っているものは正確には相関分析でなく、それは相関係数のことだ。正確に言うとピアソンさんが考えた積率相関係数だ」

佐藤「それ、本に書いてありました。でも式を見ても感じがつかめません」

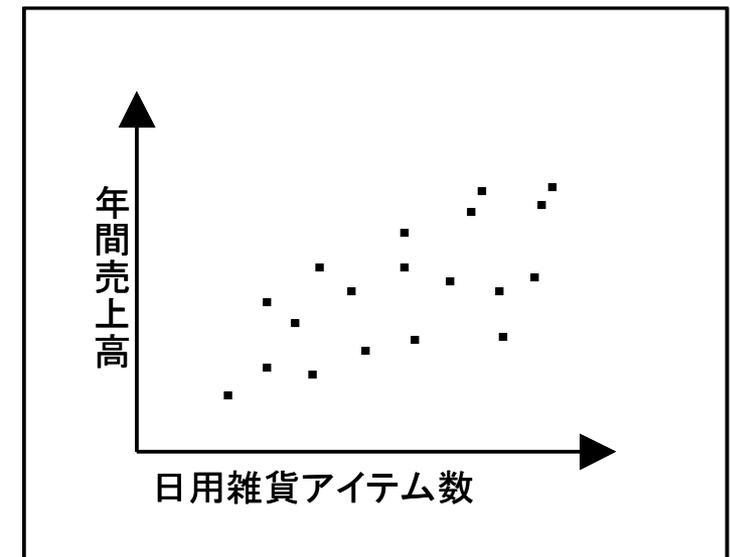
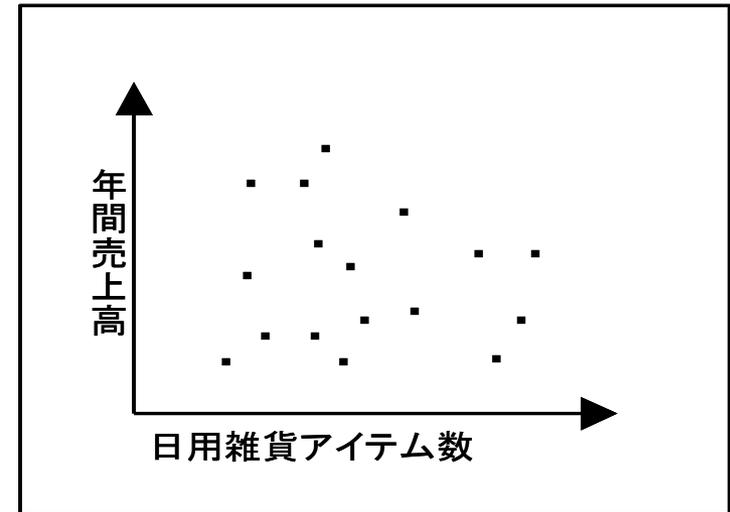
土屋「今度の売上予測の場合で考えよう。例えば他店の年間売上高と日用雑貨のアイテム*数をとってプロットしてみたら、こんな風になったとしよう」

と言って土屋は紙にグラフを書いた。(図3-7)

土屋「どう思う？」

佐藤「年間売上高と日用雑貨のアイテム数は関係ないんじゃないですか」

土屋「相関とは互いに関係するという意味だから、この時は『相関なし』だ。『なし』なのだから、何か式を使って、計算して0になったりするとそれらしいよな。もちろんこの場合でも距離の2乗の和を最小にすればいいんだから回帰直線は引けるけど、予測しても意味ないよな。つまり日用雑貨のアイテム数で年間売上高を推定すべきじゃないよな。じゃあこうなったら？」(図3-8・下)

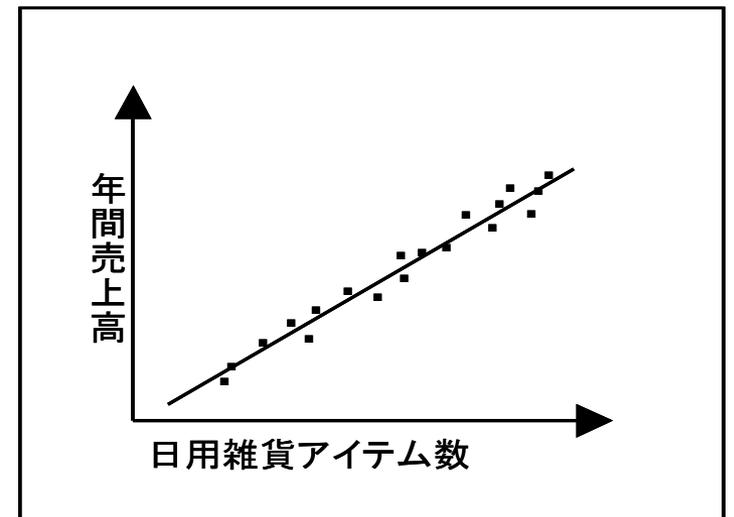
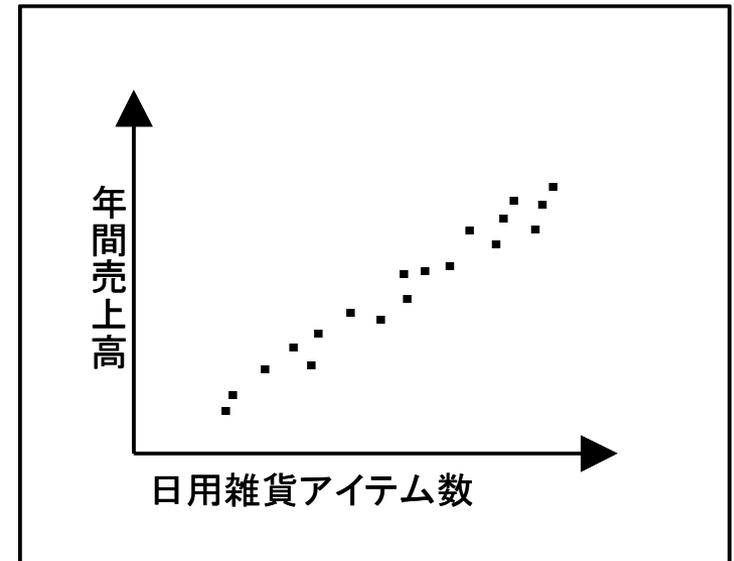


佐藤「日用雑貨のアイテム数が増えると年間売上高がそれに連れて増えてますよね。前のこづかいと年収のようですね」

土屋「これが相関ありだ。この時は日用雑貨のアイテム数で年間売上高を推定できそうだよな。これではどうだ」(図3-9・上)

佐藤「これは関係がかなりはっきりしてしてきましたね、これで回帰直線を引くと、点と線の距離の2乗の和は随分小さいですね」(下)

土屋「そうだ、これが強い相関だ。佐藤が気づいたように回帰直線と点の距離の2乗の和のようなもので、どれ位相関が強いか分かるよな。ピアソンさんは頭がいいので、それを式で計算するようにしたんだ。つまり直線と点が完全に重なったら1にして、離れたら少しずつ小さくして行って、まるっきり関係がない状態を0にすれば良いわけだ」



土屋 「次はどうだ？」(図3-11)

佐藤 「これは日用雑貨のアイテム数が少ない店の方が年間売上高が大きいんですね。こういうこともあり得るなあ」

土屋 「これが負の相関だ、片方が増えれば片方が減るというものだ。さっきとは逆の相関だ。こっちも同じように点と線が重なれば-1で、離れたら今度はだんだん大きくして行って0になれば良いわけだ。

これがピアソンが考えた相関係数だ。相関係数とは人間の相関というイメージを1から-1の数字で表したんだ。数学が物理や化学と違うのはここなんだ。物理や化学は世の中の現象を理論的に解明していくものだよな。我々ビジネスパーソンが使うビジネス数学はそうではなくて、人間が持っている感覚、直感を数字で表すんだ。強い、弱いや薄い、濃い、あるいは起きそう、起きそうもないといった感覚を数字で表現するんだ」

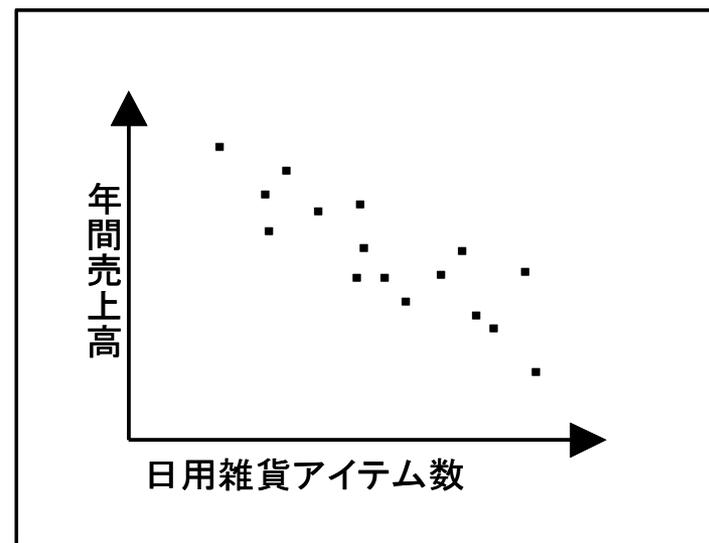


図3-11

土屋 「ところで回帰分析の説明変数にはどのデータを使えばいいと思う？」

佐藤 「相関係数の大きいもの、つまり 1 に近いものを使えばいいんでしょうか」

土屋 「- 1 ではだめか？」

佐藤 「負の相関でも使えそうですね。ということは説明変数は 1 か - 1 に近いものを選べばいいんだ。ところでどうやって計算するんですか？」

土屋 「それはコンピュータだ。さっきのややこしい式は、もう表計算ソフトが覚えてくれているから数字を入れればいい」(図3-12)

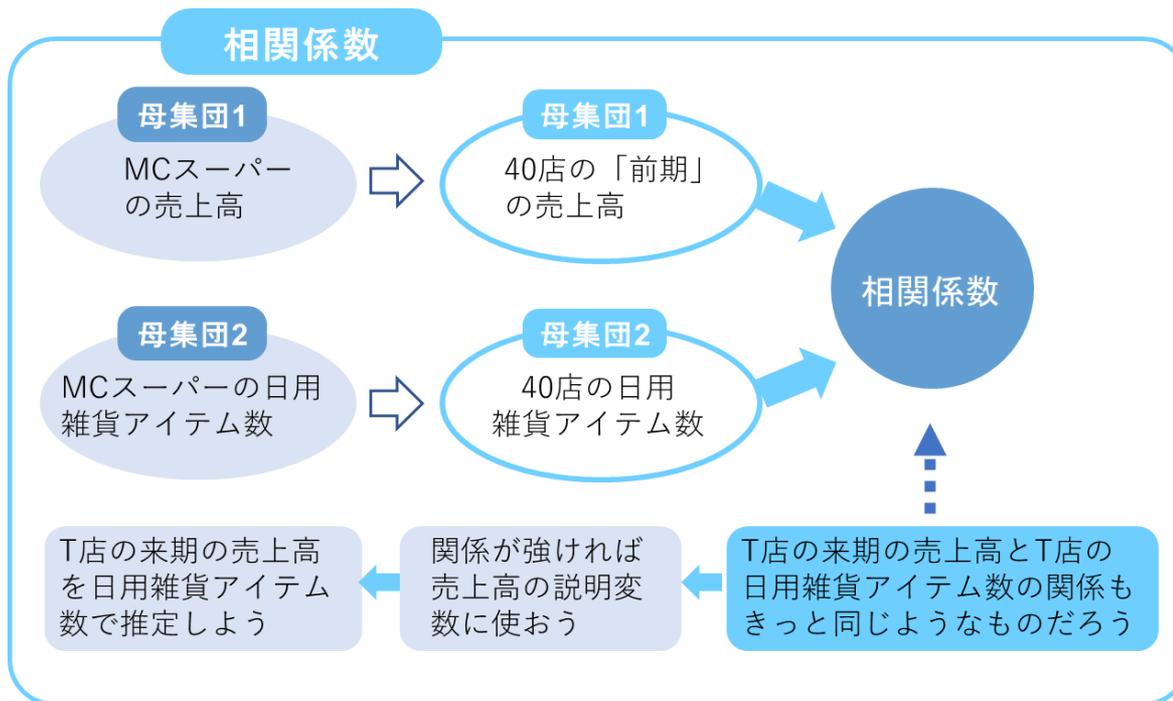


図3-12

佐藤は顧客から入手した各店舗の売上と属性データを表計算ソフトに入れた。その上で売上と各属性データとの間の相関係数を表計算ソフトで計算させた。結果は以下の通りになった。(図3-13)

前期売上高との相関係数

店舗面積	0.67
商圏人口	0.60
商圏内20～60代女性人口	0.48
商圏内世帯数	0.73
商品アイテム総数	0.66
生鮮アイテム数	0.72
加工食品アイテム数	0.51
日用雑貨アイテム数	0.18

図3-13

佐藤 「土屋さんまいりました、皆似たようなものですよ。ただ日用雑貨のアイテム数はあまり関係ないようです。でも意外と20～60代の女性人口って売上に関係ないんですね。そうかもしれないな。主婦が買い物するんだらうけど、買っていくものは家庭で食べるんだから、売上の大きさは世帯数の方が大きく影響するかもしれないな」

土屋 「別に説明変数は1つにしなくても回帰分析はできるよ」

佐藤 「でも日用雑貨はとった方が良いでしょうね。あと残り全部だと多すぎないですか？」

土屋 「あまり説明変数が多いとお客様が予測結果を見た時、直感的なイメージがわきづらいよ。何が予測した売上の根拠かとらえづらくなるんだ。絞った方が良いでしょう」

佐藤 「どうやって絞るんですか？相関係数いくつ以上は適用できるというのがあるんですか？」

土屋 「そんなのはないよ。でも説明変数同士に関連のあるものは、相関係数が高いものを1つ選んだ方がわかりやすく、お客様に説明しやすいと思うよ」

佐藤 「説明しやすいなんて理由で選んで良いんですか」

土屋「良いんだ。例えば仮にこの店舗売上高と商圈内にある信号機の数がかかなり高い相関を示したとしよう。この時信号機の数を使ってT店の売上高を予測したらお客様は何ていうと思う。『売上と信号機の相関が高いなんて、たまたまそうなったんじゃないの』と言うだろう。その予測売上高があたっていると、説得力がない予測なんて、直感がわからない予測なんて意味ないだろう。もっと言えば店舗売上高と相関をとることのできるデータなんて無限にあり、そのデータ同士も相関があったり、なかったりだろう。そう考えればもっとも予測精度の高い説明変数を選ぶのでなく、『佐藤がお客様に数字を説明する』という目的にもっともあった、佐藤もお客様も納得できる直感のわく説明変数を選ぶことが大切なんだ」

佐藤「そういう目で見るとこのデータは、店舗面積と人口・世帯数とアイテム数の3つのグループに分けることができます。それぞれのグループの中から相関係数の大きいものをとろう。そうすると、店舗面積と商圈内世帯数と生鮮アイテム数の3つになるな。世帯数には人口が加味されているし、商品アイテム総数は店舗面積に加味されているから、良い感じかなあ」

土屋 「このように説明変数が2つ以上になるものを重回帰分析というんだ。こづかいと年収のように説明変数が1つのものは単回帰分析だ。よし、ここまでわかったら回帰分析についてもう一度しっかり勉強しろ」

佐藤 「そんなこといわないで今教えて下さいよ」

土屋 「まず自分で考えろ。数字を使うには、耳でなく頭を使うことだよ」

佐藤は「誰でもわかる統計の本」の回帰分析の続きを読み始めた。(図3-14)

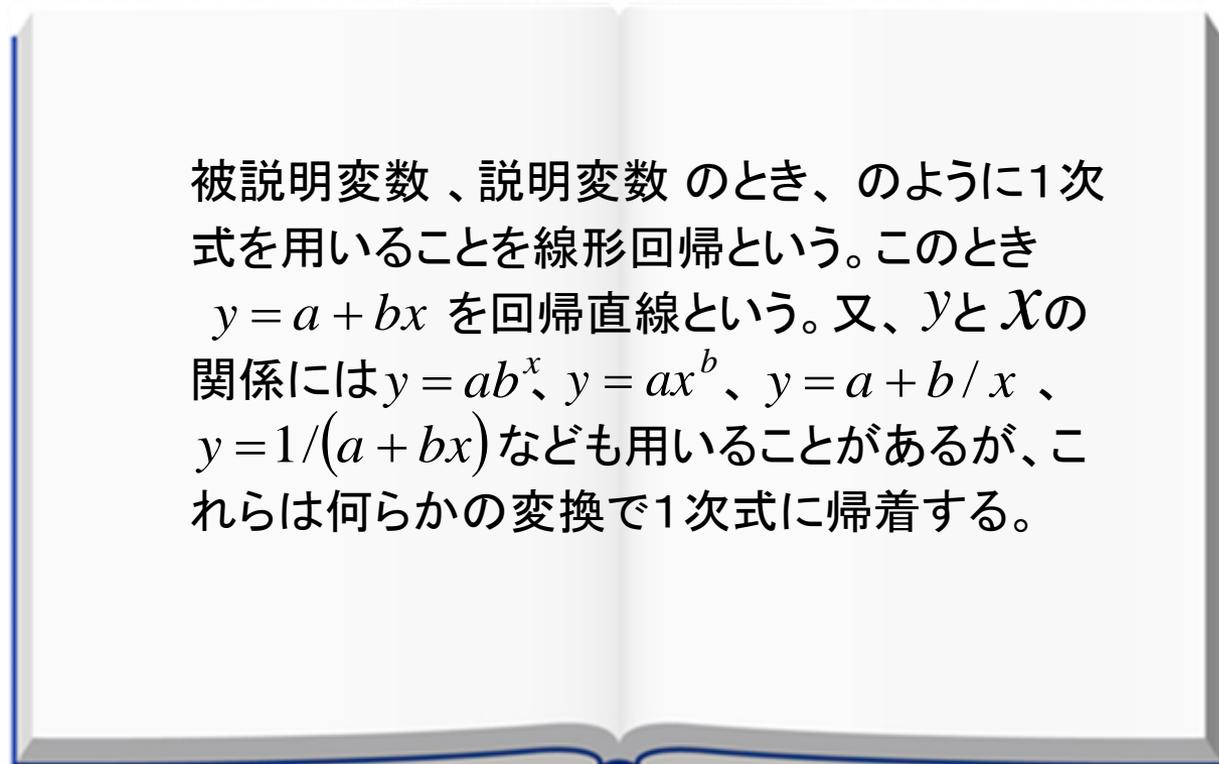


図3-14

直線の式はこづかいと年収でやったやつだな。直線だから線形か。この後の不思議な式は何だ。何らかの変換って何だ

翌朝土屋に質問した。

佐藤「この式は何ですか？」

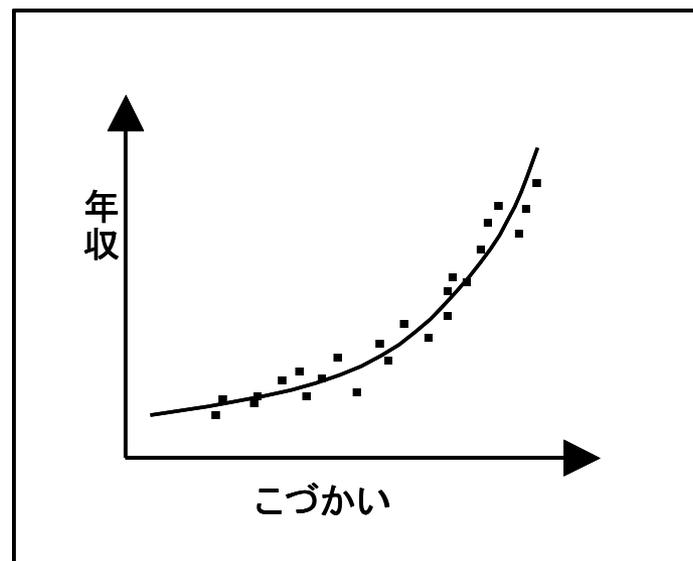
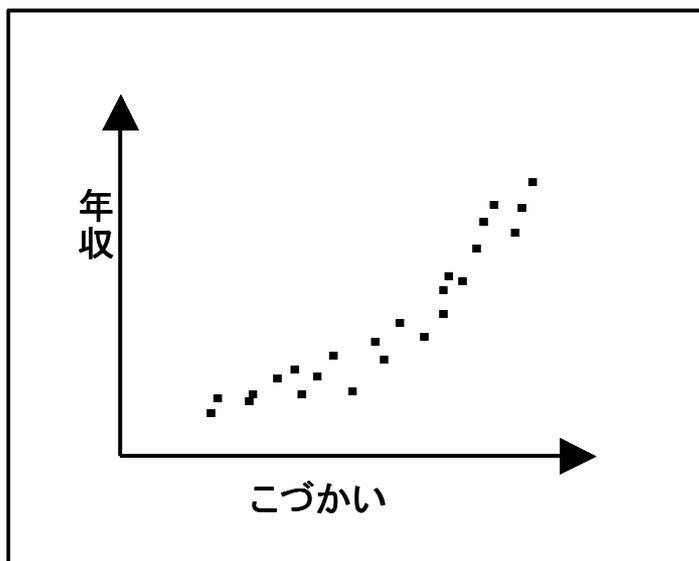
土屋「これは被説明変数と説明変数の関係が直線で表しづらい時に使うんだ。たとえばこづかいと年収をプロットしたらこんな風になったとしよう」

と土屋は紙にさっとプロットした。(図3-15・下左図)

土屋「どんな線が見えるか？」

佐藤「直線で表せないこともないですが、こんな感じの線ですかね」

佐藤は土屋が書いた点の上に曲線を引いた。(図3-16・下右図)



土屋「そう思ったらそんな曲線を表す式を使えばいい。こんな時はどうだ」

(図3-17)

佐藤「こんな感じかな」 (図3-18)

土屋「そのときはこれにあてはまる式を選べば良いんだ。曲線にもいくつかのパターンがあるので、その本にもたくさんの式が書いてあるんだ」

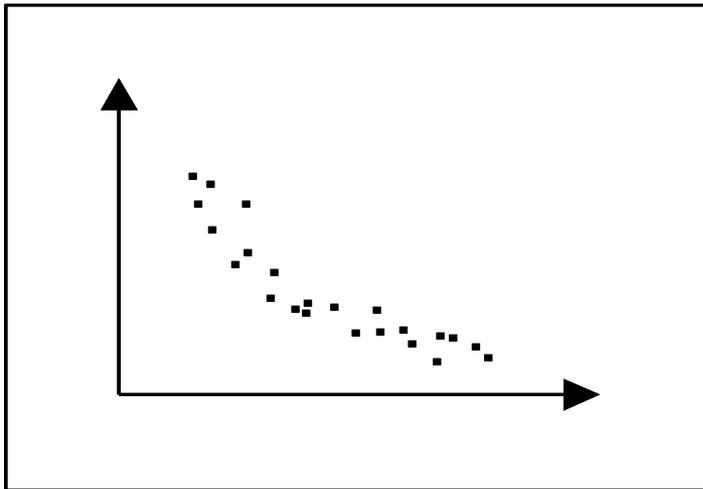


図3-17

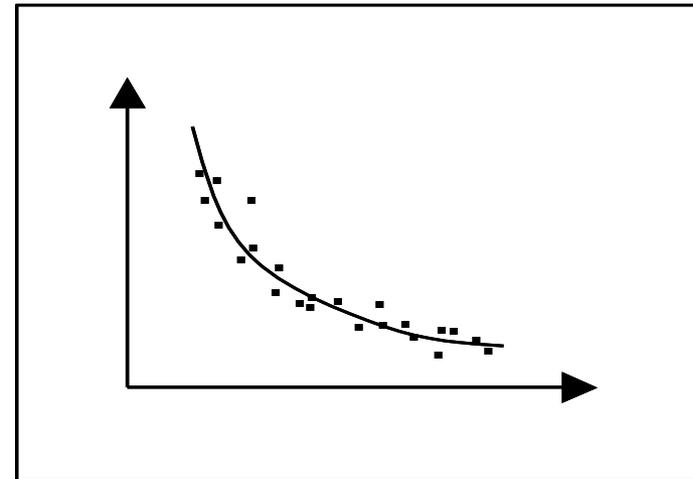


図3-18

佐藤 「そういえば中学の頃に曲線の式もやったような気がします」

土屋 「まず式を選ぶんじゃなくて、プロットしてみて、自分のイメージを大切に
して、線を引き、その線に合った式を本で探せばいいんだ。それが前にも
言ったニュートンの万有引力の法則とちがうところだ。データから考えて人
間のイメージに合った式を選び、またデータがたまってきたら、状況が変わっ
てきたら、違う式を使えばいい」

佐藤 「ところでこの本に書いてある何らかの変換ってどういう意味ですか？」

土屋 「それは、佐藤が学校でやったけど忘れてるか、やらなかったかもしれない対数などを使えば、曲線の式も直線の式に変えることができるということ言ってるんだ。つまり曲線で式を表しても直線のやり方を知っていればできる、ということ言っているんだ」

佐藤 「今回は3つも説明変数があるので、どこかで曲線を使うことになりそうですね」

土屋 「心配するな。今回のようにMCスーパーの店舗という限られた、そして似たようなデータを集めて行う時はすべて直線で考えて問題ない。例えばさっきのこづかいと年収を曲線で表したとしても、一部の似たようなデータだけ見れば直線でも問題ないはずだ。日本全国のビジネスパーソンでは曲線であっても、ミドリ食品の従業員だけなら直線でOKのはずだ。ミドリ食品の従業員なら大体似たような年収で、似たようなこづかいだろう。まさかこづかい月50万円で年収5000万円の人なんていないだろう」

とって土屋は先ほどのこづかいと年収のグラフの一部を四角で囲った。

([図3-19](#))

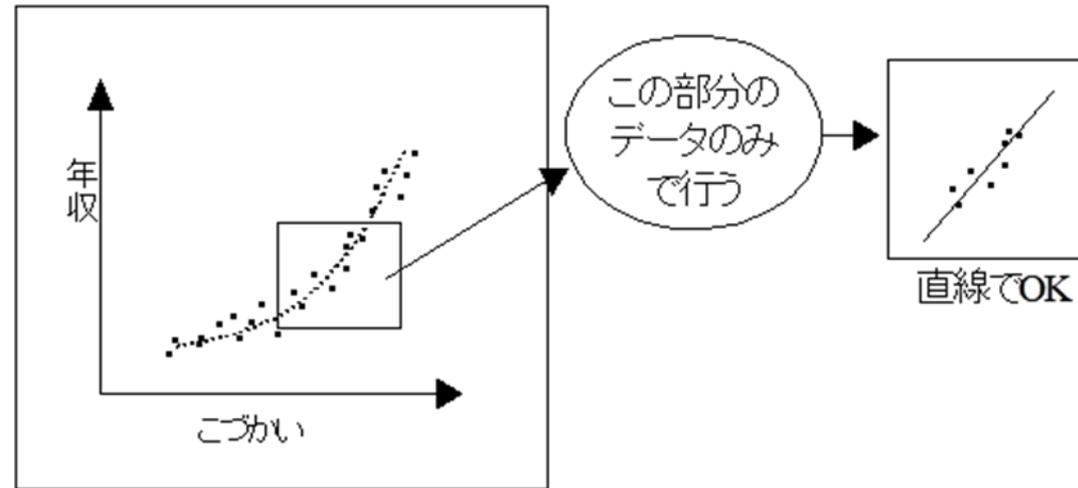


図3-19

土屋 「この四角の部分を見てみる、これだけなら直感では直線を引くだろう。まあ佐藤がもっとえらくなって、うちのマーケティング本部長かなんかになって、MCスーパーだけでなく日本中のスーパーマーケットの分析をしなくてはならなくなった時は、ちょっと話が違うが」

佐藤 「よかった。今回はすべて直線で考えれば良いんですね。もう一度本を読んで考えてみます」

佐藤は帰宅して「誰でもわかる統計の本」の回帰分析のつづきを読んでいった。
(図3-20)

説明変数が1つのものを単回帰分析、複数あるものを重回帰分析という。重回帰分析では説明変数を x_1 、 x_2 、 $x_3 \cdots$ としたとき

$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 \cdots + a_nx_n$$

と表されるものを加法モデル、

$$y = a_0x_1^{a1} \cdot x_2^{a2} \cdots x_n^{an}$$

と表されるものを乗法モデルという

乗法モデルは対数変換によって加法モデルに帰着する

図3-20

「前半の加法モデルというやつが直線で、後半の乗法モデルというのは曲線だな。曲線は無視だ。すべて直線だ。だけど説明変数がたくさんあるのに、どうやって直線を引くんだろう」

翌日、佐藤はこれを質問した。

土屋「まず説明変数2つの時で考えよう。この時は3次元空間だ」

佐藤「3次元ってなんですか」

土屋「前に x と y で書いたときは『横とたて』だったから、横とたてで2次元だ。これに高さを足すんだ。これで3次元だ。いってみれば正方形と立方体のちがいだ。紙に書いてあげよう。横が x_1 、たてが x_2 、高さが y だ。例えば x_1 が店舗面積、 x_2 が商圏内世帯数という説明変数で、 y が売上高という被説明変数だ。この立方体の中に前と同じように点がプロットされていくんだ。立方体の表面じゃないぞ、中の空間に点があるんだ」

とって紙に立方体の絵を書いた。(図3-21)

土屋 「イメージできたか」

佐藤 「要するに大きな立方体の箱の中に点が浮かんでいると思えば良いのですね。空中に浮かんだゴミのようなものですね」

土屋 「そこに前と同じようにすべての点からの距離がもっとも小さくなるように直線を引けば良いんだ」

佐藤 「うまく引けません」

土屋 「これもコンピュータの仕事だ。前に説明した最小2乗法と微分だ。つまり佐藤はわからなくていい世界だ。佐藤はただ距離の和を最小にするように線を引いていると思えばOKだ。そうすれば前にやった平面のときの回帰分析と同じだろう」

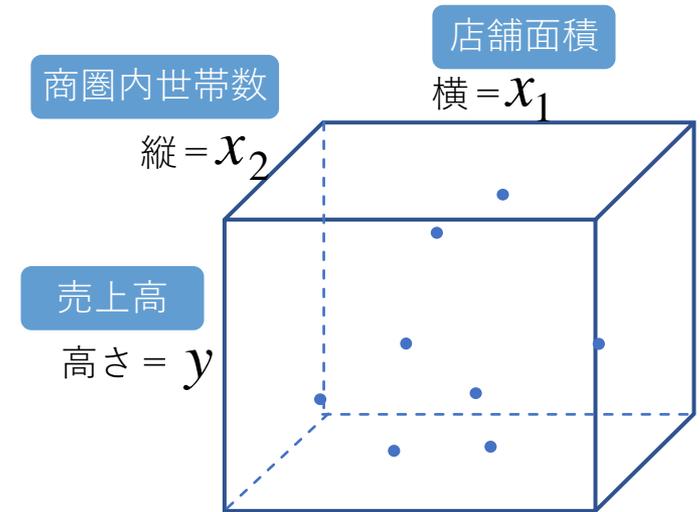


図3-21

佐藤 「わかりました。というより説明変数が1つの時も、ここの所の細かいやり方はどうせわからなかったんです。でもMCスーパーの場合はもう1つの生鮮アイテム数という説明変数があって、全部で3つあるんですけど大変ですよ、2つの説明変数で3次元だから、3つだと4次元の世界になってしまいますよね。だんだんミステリアスになってきましたね」

土屋 「確かにそうだが、横軸、たて軸に、高さという軸を加えることができたんだから、もう1つ位何か軸が増えたって線は引けるだろう」

佐藤 「うーん、引けません」

土屋 「要するにたて、横、高さ以外にもう1つ何か要素があって、その中に点があり、その点からの距離を最小にすれば良いんだろう」

佐藤 「わかったような、わからないような」

土屋 「そういうものなんだ数学は。実際には頭に描けないものも考えていくんだ。ただ佐藤は数学者になるんじゃないだろう。仕事で使うんだろう。これを理解する目的は何だ」

佐藤 「お客様にきちんと自分の出した数字の意味を説明できることです」

土屋 「だったら細かい計算法などどうでも良くて、その考え方を理解することだろう。単回帰分析の延長に重回帰分析があることを3次元の例でわかれば十分だろう。もう一度こづかいと年収を思い出せ。佐藤は回帰分析をどういうふうに理解した？」

佐藤 「年収を知りたいがわからないので、代わりにこづかいを調べて年収を推定する。そのためにはこづかいと年収の両方がわかっている人のデータをできるだけ集めて、そのデータからこづかいと年収の関係をパターン化しておく。難しくいうと回帰してですね、そしてその回帰式を使ってこづかいから年収をあてるんですね」

土屋「よく理解できたじゃないか。ただ最後の『あてる』が気になるなあ。年収をあてるんなら直感の方がいいかもしれないぞ、そこは『今あるデータから考えてみると、その人の年収はこの金額とすることが妥当だと考えられます』ということだろう。これを今回は、年収をT店の売上高、こづかいを店舗面積、世帯数、生鮮アイテム数に置き換えれば良いんだ。ここまでわかればあとは表計算ソフトに数字を入れて回帰するだけだ」

佐藤はMCスーパーの木村氏にプレゼンテーションを行っていた。

佐藤 「いただきました他店舗データをもとに重回帰分析にて、T店の年間売上高を推定しました。説明変数としては店舗面積、商圈内世帯数、生鮮アイテム数を用いました。その結果は約8億8000万円となりました」

木村 「重回帰分析って何？」

佐藤 「要するに他店舗のデータから見て、T店の売上をどれ位と考えるのが妥当かと考えるものです。いってみれば各店舗の平均みたいなものですが、単純に40店舗の平均ではいかにも雑で、本部長から店舗面積のちがう店の平均なんて意味がないだろうといわれてしまいますよね。そこで店舗面積、世帯数、生鮮アイテム数の3つを考慮して平均を出したようなものです」

木村 「なんでその3つなの？もっとあるんじゃないの」

佐藤 「いただいた各店舗データと売上高の相関係数を出してみました。あまり平均に考慮するものが多いと、かえってわかりづらいので、店舗面積などが売上に影響を与える度合いを相関係数という形にして計算してみました。つまり店舗面積が大きくなるとどの位売上が伸びるかということです。その結果はお手元の資料のとおりです。商圈人口、世帯数、20～60代女性人口は互いに関係が強いので、1つにしようと思い、もっとも相関係数の大きい世帯数を選びました。取扱商品のアイテム数についてはやはりもっとも相関係数の大きい生鮮アイテム数にしてみました」

木村 「でも生鮮アイテム数が多いほど売上が大きいというなら、店の中をすべて生鮮にした方が良いということにならないか」

佐藤 「いただいた店舗とまるで違うタイプの店舗は、このデータでは予測することができません。難しくいえば生鮮だけの店舗は今回の標本とは母集団が違ふと考えるべきです。T店は今回いただいたような40店舗と似たような、つまり生鮮、加工食品、日用雑貨の取扱アイテム数のバランスがある程度似ているというのが前提です」

木村 「ところで後になって思ったんだが、1つ大事なことを忘れていた。競合店の影響を考えないわけにはいかないと思うんだ。近くに同じような食品スーパー*や大型総合スーパー*があるんじゃ、かなり売上が違ってくるよなあ」

佐藤 「わかりました。ところで今回の40店舗については競合店のデータはあるんですか？」

木村 「今度のT店を含めて持っているよ。ただ食品スーパーと大型総合スーパーがそれぞれ商圈内にあるか、ないかだけだけどね」

食品スーパー、総合スーパー

食料品を中心とした品揃えでセルフサービスによる大量販売を行う小売業をスーパーマーケットと言う。スーパーマーケットのうち、品揃えを衣食住のすべてに広げたものを総合スーパーと言う。総合スーパーと区別する意味で前記のスーパーマーケットを食品スーパーとも言う。

佐藤は心の中で「まいった」と思った。「食品スーパー、大型総合スーパーって数字で表せないよなあ。また土屋さんに聞くか」

佐藤「数字で表せないものって回帰分析できないですよね？」

土屋「できない。だけど数字に表せないものなんてあるのか？」

佐藤「MCスーパーから言われたのは競合店のあり、なしを考慮しろということなんです」

土屋「数字にできているじゃないか、『あり』が1で、『なし』が0でいいんだろう」

佐藤「でも食品スーパーと大型総合スーパーの2つがあるんですよ、どうやって数字にすればいいんですか？」

土屋「何度も言うように、まあ大体佐藤が悩むようなことは昔から皆悩んでいるんだから心配するな。説明変数を2つ使えばいいんだろう。食品スーパー変数は『あり』が1、『なし』が0、大型総合スーパー変数は『あり』が1、『なし』が0だ。これが数量化理論だ。数字になっていない質的データを数量化していく考え方だ、簡単だろう」

佐藤 「そうか、前の3つにこの2つの説明変数を加えればいいんだ。そう考えると変数が5個だから6次元だ。まあ4次元も6次元もどうせわからないから関係ないか。表計算ソフトに入れば同じですね」

土屋 「このようにたくさん変数を使っていろいろな分析をしていくことを多変量解析というんだ。多変量解析にはいろいろな方法が考えられており、重回帰分析もその1つだ。まあ重回帰がわかれば後のやり方も理解できるよ」

佐藤は再びMCスーパーの木村にプレゼンテーションしていた。

佐藤「食品スーパーのあり、なし、大型総合スーパーのあり、なしを数量化理論を用いて回帰分析をした結果、お手元の資料のようになりました」

木村「これって便利だなあ。前からつくづく思っていたことがあるんだ。全社売上目標を各店舗へ割り振る時、どうも声の大きい店長が勝っていたような気がする。声の小さい店長は大きな目標を押しつけられて、結局目標未達で評価が低いよなあ。自分はこれから新人店長なんで、店長会議などでとてつもなく大きい目標を押し付けられそうで不安だったんだ。これを使って配分すれば公平だよな。佐藤さんの言いたいことよくわかったよ。売上予測はあたる、あたらないじゃなく、過去のデータから考えた平均のようなものというのは新店のことを考える私にはまさに実感できるよ」

佐藤「そういえばMCスーパーさんはポイントカード*を出していましたよね」

木村「ああ出してるよ。ポイントカード発行時には入会書をお客様に書いてもらっているよ。これでお客様の住所、氏名だけでなく年齢、性、職業、来店手段や家族構成までわかるんだ」

佐藤「お客様ごとの購入金額もわかるんですよね」

木村「100円で1ポイント出してるから、ポイントでわかるよ」

佐藤「それを使えば、いろいろ分析できますね」

木村「今まで年令階層別購買金額表とか、来店手段別購買金額表といった色々なレポートを本部からもらっていたけど、何かよく使い方がわからなかったんだ。」

佐藤「この回帰分析を使えば購買金額を被説明変数、つまりさっきの店舗売上高として、カード記入項目を説明変数、つまりさっきの店舗面積などにすれば購買金額の式が出せますよ」

木村「うちの社長はよく『店に来るお客を購買金額で分けてみる、高い人がカスタマーで、低い人がゲストだ。次にカスタマーがどういう人かを分析しろ』とっている。そうかこの式をベースにカスタマーの基準を決めて、分析していけば良いんだ」

佐藤「というよりも、それを使って、その平均的購買金額を出していないお客様をさがすのはどうですか？40代の専業主婦で、家族構成が5人で、徒歩で来店している最高のはずのカスタマーが、回帰式から出した予測金額に行っていない人なんかをさがすんですよ。この人たちはきっと他店で買っているはずですよ。この人たちがこちらを向くような企画や、商品の品揃えに持っていけばいいんじゃないですか」

木村「それがうちの本部長が知っているワン・トゥ・ワン・マーケティング*だ。
新店オープンしたら、佐藤さん一緒にやってみようよ」

ポイントカード

小売業、サービス業において一定の購買金額に応じてポイントを発行し、そのポイントに応じて値引きなどさまざまなサービスを行うことをポイントサービスと言う。このポイントを蓄積・表示するカードをポイントカードと言う。

ワン・トゥ・ワン・マーケティング

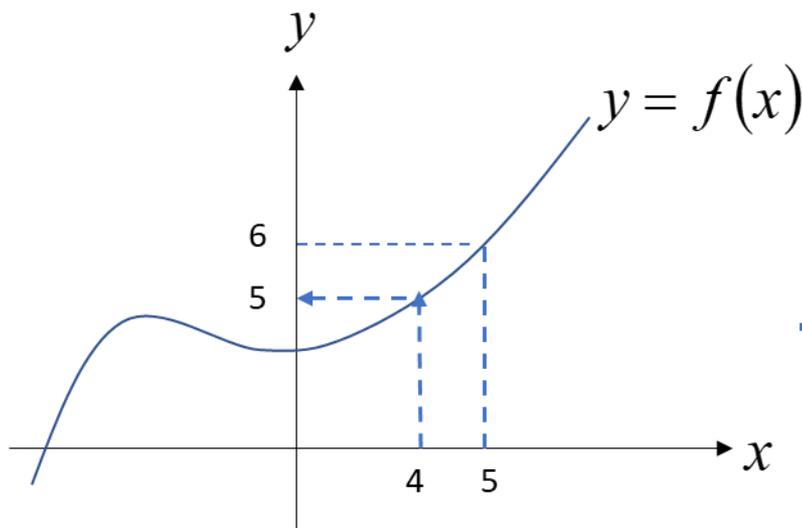
顧客一人ひとりに対応してプロモーションなどマーケティングのやり方を変えていくことを言う。

指数と対数

変数 (parameter: 日本語でもパラメータともいう) とは、言葉のとおり数値が変わっていくものをいいます。代表的な変数が「時間」です。まさに時々刻々と変わります。一方、変わらない数値を定数 (constant: よく C で表す) といいます。正確にいうと、定数とは、何かを考えると、「変わらない」と考えるものです。絶対に変わらないものは不変です。

変数が2つあって一方が決まると、一方が決まるような「関係」を関数 (function: よく f で表す) といいます。変数、 x があって、 y が決まると y が決まるとき、 $y = f(x)$ と表します。

$y = f(x)$ は x を横軸、 y をたて軸としたグラフでよく表現されます。

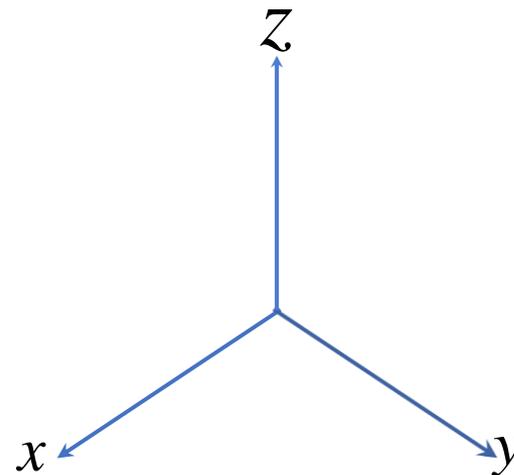


→ x が4と決まると y は5、
 x が5と決まると y は6

x 、 y という文字を使ったのは、アルファベットの終わりの3つである x 、 y 、 z を使って3次元を表そうとしていたようです。そのうち人間が3次元についていけず、紙にきちんと表現できる2次元の x 、 y で表されるようになりました。

例えば、MCスーパーが無人の自動販売機だけの小さな店舗を開発し、とりあえず1店舗出しました。社長はこれをこの後、毎年倍増していくと決めたとします。そうすると、1年後は2店舗、2年後は4店舗…となり年 (x 年) が決まると店舗数 (y 店) が決まりますので、 x と y は関数といえます。

これは「 $y = 2^x$ 」と表されます。 2^2 は 2×2 、 2^3 は $2 \times 2 \times 2$ です。これを指数関数といいます。



社長は「ところで100店舗になるには何年かかるんだろう」と考えました。今度は店舗数(X 店)を決めて年数(Y 年)を求めるわけです。 $2 \times 2 \times 2 \times \dots$ とやっていって、7回かけると100を超えて128店舗となります。つまり7年後です。

この時、 $\log_2 128 = 7$ と書きます。 $\log_2 128 = 7$ とは「2を何回かけると128」になるかということです。またこの2を底といいます。これを X (店舗数)、 Y (年数) の関数として考えると、 $y = \log_2 x$ となります。これを対数関数といいます。

データを表計算ソフトなどを使ってグラフにしようとするとうるんことがあります。小さい数と大きい数がある時です。10、100、1000、10,000という数字を1つのグラフにしようすると、10に合わせると10,000ははみ出てしまいますし、10,000にあわせるとすると10が小さすぎて見えなくなります。この時、対数を使うと便利です。10を底とする対数を考えれば $\log_{10} 10 = 1$ 、 $\log_{10} 100 = 2$ 、 $\log_{10} 1000 = 3$ 、 $\log_{10} 10,000 = 4$ となり1、2、3、4という数字になります。対数目盛りのグラフとはこの対数で表していくグラフのことです。

また対数には便利な性質があつて

$$\log_a (x \times y) = \log_a x + \log_a y$$

となります。

ところで $y = ab^x$ (本文[153ページ](#))という関数を回帰式として使いたくなつたとします。これは底 b の対数をとると $\log_b y = \log_b a + x$ となります。

ここで $\log_b y = z$ 、 $\log_b a = c$ とおくと $z = c + x$ となり、 x 、 y という2つの変数の関係は直線の式で表され、直線と考えて扱えるようになります。

これが本文[153ページ](#)に書いてあった「直線に帰着する」という意味です。

第 4 章

発注ロットを考える

ミドリ食品では小売店からの注文の小口化に系列の卸*とともに苦慮していた。コンビニエンスストアから生まれた多頻度少量発注、小口配送*の波はスーパーマーケット、一般店にも押し寄せており、在庫増大、物流コストのアップに頭を痛めていた。

中村支店長「佐藤、ちょっと来てくれ。今私は毎月本社でやっている取引適正化委員会に出ているんだ。その委員会で各支店で1人ずつ担当を決めて、委員会と現場のパイプ役をやってもらうことになった。この支店では君がやってくれ。今、小売との取引がどうなっているかよく整理し、会社として今後どうすれば良いかを私と一緒に考えていって欲しい。当面はMCスーパーを中心に考えてくれればOKだ。いきなりMCスーパーと話すんじゃなく、まず今一番問題となっているコンビニの発注方法についてよく調べ、今の小売の動きを理解してから取り組んでくれ」

 系列の卸

製品メーカーは自社製品の販売のため、自社製品専売の卸売会社を作ったり、契約したりすることも多い。これを系列卸と言う。

多頻度少量発注、小口配送

小売業において、発注のサイクルを短くし、商品を少しずつ配送してもらうことを言う。これによって店舗での在庫量が削減され、さまざまな商品を置くことが可能になる。

しかしミドリ食品ではコンビニエンスストアとの取引はすべて本社の営業本部で行っており、S支店にはその担当がいなかった。

「発注方法を調べるなら、発注のプロに聞けばいいはずだ。えーと、そうだ、この支店の総務課の山田さんが、前に本社の資材部にいたはずだから聞いてみよう」

佐藤「山田さん、コンビニエンスストアの発注方法がどうなっているか知りませんか」

山田「そういえば前に勉強したことがあるよ。確かトヨタのカンバン方式をマネして作ったやり方だと聞いたなあ」

佐藤「カンバン方式って何ですか？」

山田「欲しいものを、欲しい時に、欲しい量だけ買うことだ。そう考えていくと発注ロット*は、どんどん小さくなっていくんだ」

佐藤「欲しいものを買うのは当然でしょう。いらないものを買ってもしようがないでしょう。でもどうして、そうすると発注ロットは小さくなるんですか」

山田「今君と話してもわかってもらえそうもないね。発注について知りたいなら、まず基本の基本ともいえる経済的発注量について勉強してから来てくれ」



発注ロット

1回の注文で買う商品の最低個数。「何個ずつ買うか」ということ。

佐藤は本屋で経済的発注量のことを書いてある「誰でもわかる在庫管理のやり方」という本を買った。

経済的発注量

需要がほぼ安定している場合に、1回当たりの発注ロットを大きくすると発注費用は低くなるが、在庫費用は高くなる。反対に、発注ロットを小さくすると発注費用は高くなり、在庫費用は低くすむ。この2つのトレードオフの関係のなかで、最も総費用が低くなるロットサイズを経済的発注量(EOQ: Economic Order Quantity)といい、次の式で計算される。

$$EOQ = \sqrt{\frac{2 \times \text{1回あたり発注費用} \times \text{総需要量}}{\text{1個あたり年間在庫費用}}}$$

図4-1

「書いてあることは何となくはわかるが、まず式がわからない。そうだと土屋さんに言われたことを思い出そう。式を見るのでなく、意味を理解しろだ。要するに商品などを発注すると2つの費用がかかるということだ。1つは在庫費用でもう1つは発注費用か。在庫費用はその商品を保管する費用だろう。発注費用って電話代か？ どうしてトレードオフなんだ？」

翌朝山田に聞いた。

佐藤 「在庫費用って保管する費用のことですよね」

山田 「大体あってるよ。ただ保管する倉庫代や店舗での陳列の場所代の他にも、売れなくなった時の廃棄ロス※なんかも入るね、そんなのを全部含めてだ」

佐藤 「発注ロットを増やすと、店に届く商品の量が増えるので、在庫費用がかかるんですよ」

山田 「そうだと、発注ロットに在庫費用は比例すると考えて良い。つまり発注ロットが2倍になれば、在庫費用も2倍になるということだ」

佐藤 「10個ずつ買っていたものを20個ずつにすると、2倍の在庫費用か、何となくわかる。ところで発注費用って電話代のことですか」

山田 「発注は今はあまり電話を使わずコンピュータでやってるのは知っているだろう。ただ発注費用の中にはそういう通信費もあるけど、そんなものは大したことはない」

佐藤 「じゃあ、何ですか」

山田 「物流費用だ。配送費用といってもいい」

佐藤 「そうか発注ロットを大きくすると発注回数が減って、店に行く配送の回数も減るし、小さくすると配送回数が増えるということか、10個ずつ買っていたものを20個ずつにすると、同じだけ売れるなら2回に1回行けば良いわけだ」

山田 「そうだ、反比例だ。片方増やすと片方減るというやつだ」

佐藤 「それって、負の相関に似てますよね」

山田 「あれ、よく知っているね、その通りだ。そう考えると2つの費用はこんな線になって、その2つの費用の和はこんな風な線になるだろう」

山田は紙にグラフを書いた。(図4-2)

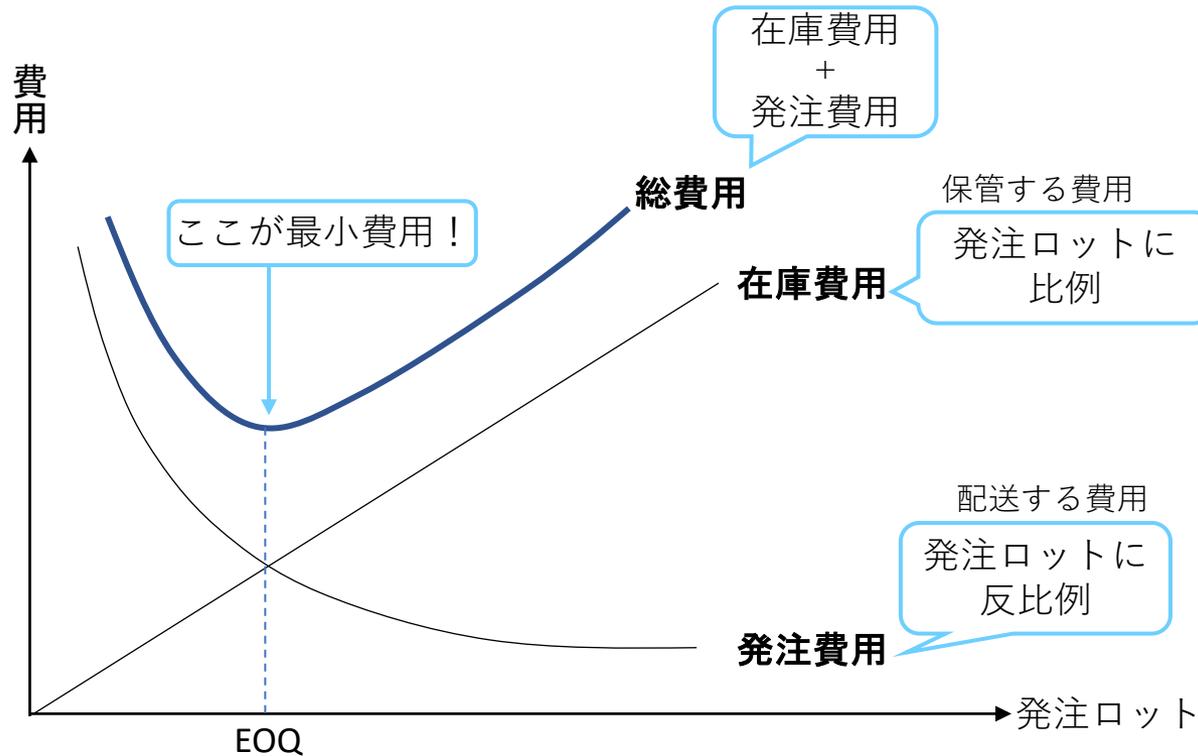


図4-2

山田 「在庫費用は発注ロットが2倍で2倍となるから、こういう直線だ。発注費用は発注ロットが2倍になれば半分だからこんなグラフだ。双曲線だ。中学でやったろう。その2つの費用の和はこんな曲線になるんだ」

佐藤 「ということは最小の費用があるんですね」

山田 「そうだ、その点がEOQ、経済的発注量だ。総費用を最小にする発注ロットの大きさだ」

佐藤 「もしかして在庫費用と発注費用を足して微分するんでしょう」

山田 「よくわかっているな」

佐藤 「最小値を出す時は常識でしょう。そうか、ということはこのEOQを発注ロットに使えるば費用が最小になるんだ。じゃあ、なぜコンビニは発注ロットをどんどん小さくしていったんですか？」

山田 「発注ロットを小さくすると、つまりグラフを左側に持っていくとどうなる？」

佐藤 「発注費用、つまり配送費用が加速度的に大きくなってしまいますよね」

山田 「だからうちは困ってるんじゃないか。ただコンビニ側から見てみる、配送費用はどこが負担してる？」

佐藤 「そうか、普通は配送費用は売る側、つまりメーカーや卸が負担しているんだ。商品の納入単価は1回に何個ずつ買うかというよりも、期のトータルで何個買ってくれるかで決めるものなあ」

山田 「もし発注ロットによって納入価格を変えないとすると、コンビニから見たEOQはどうなる。グラフを書いてみる」

佐藤 「在庫費用はさっきと変わらないよな。発注費用は配送費用に関係ないし、コンビニはコンピュータでやってるから、通信費用も発注ロットや発注回数にはほとんど関係ないなあ。しかも発注が増えてもアルバイトが手の空いている時にやるので、発注ロットによらず一定と考えていい、そうなると総費用はこうなる、と」(図4-3)

佐藤 「あれEOQはゼロだ」

山田 「これが組立メーカーが考えたカンバンだ。在庫費用をゼロにする。つまり部品などを必要になるまで買わず、在庫を持たないことだ」

佐藤 「でも小売店では在庫のない、つまり店に置いてない商品なんて誰も買いませんよ」

山田 「それなら1個だ。その1個が売れた瞬間にもう1個持ってくるというのが理想のスタイルだ」

佐藤 「これがコンビニの発注や、カンバンか。なるほどこれが小口配送の意味か」

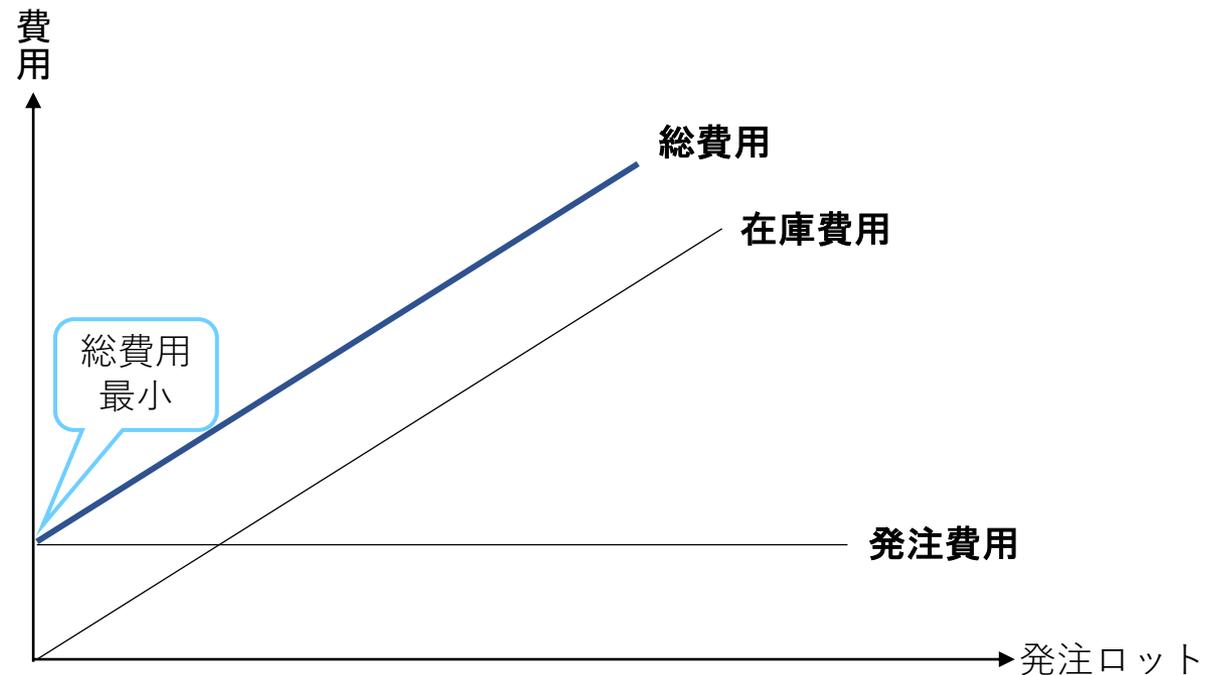


図4-3

山田「取引は常に『売る』と『買う』だ。コンビニが買う方なら、我々メーカー側はそこに売る方だ。次は売る方からこのロットサイズを考えよう。グラフを書いてみる」

佐藤「発注ロットが増えると、うちの在庫費用は減りますよね。作ったものを全部買ってくれれば、うちは在庫がないから費用はゼロだ。まあ直線でいいや、一方発注費用は配送費用と考えて、これをうちが負担すると、さっきと同じだ。だけど作った以上は売れないから、さっきの在庫ゼロの所で止めておこう。総費用はこんな感じだ」([図4-4](#))

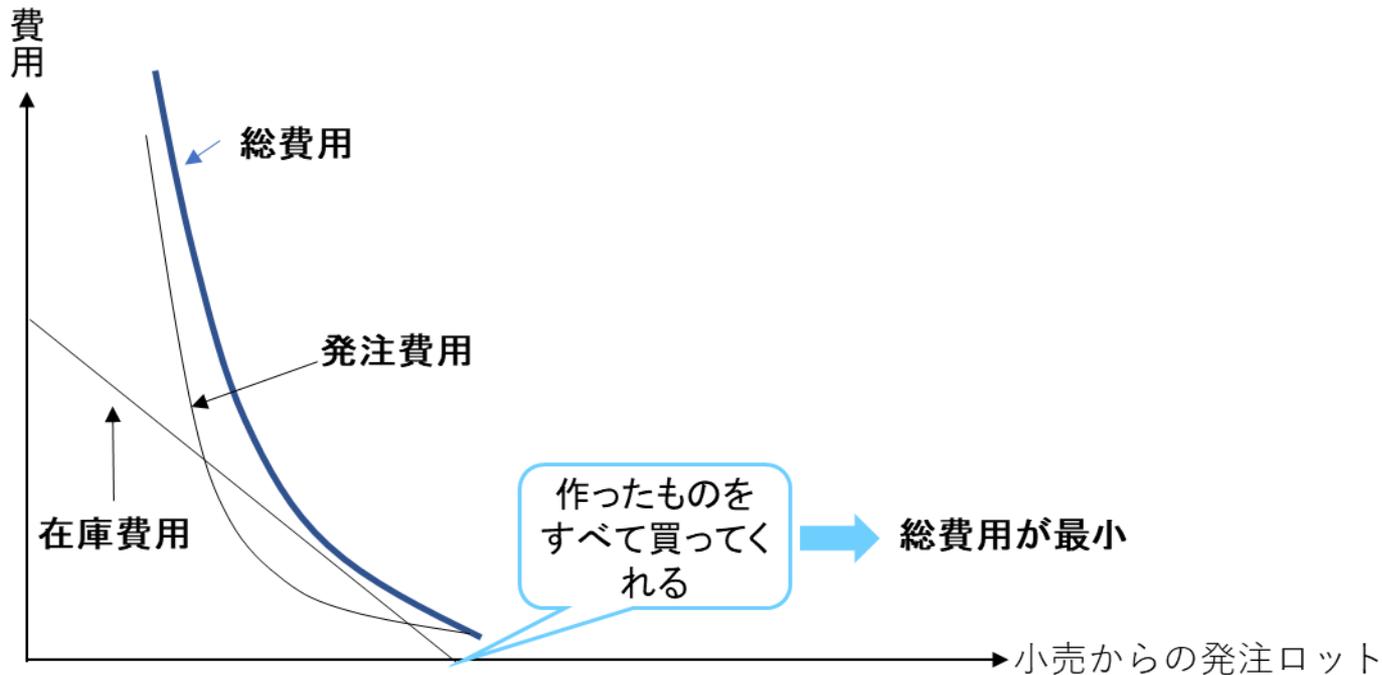


図4-4

佐藤 「ということはEOQはこの在庫がゼロの所だ。要するに大きいほど良いんで、うちが作ったものを小売が1度に全部買ってくれるのが最高だ。しかしこれじゃあ小売とロットの交渉なんてできませんよね。小売は小さいほど良く、うちは大きいほど良いんだもの。声の大きさや力関係で決まっちゃいますよね」

山田 「今、うちでは取引先の資材メーカーとS C Mというのをやってるんだけど、営業ではお客様とやってないのか？」

佐藤 「やろうとしています」

山田 「S C Mって売る方と買う方が一緒に幸せになることを考えるんだらう」

佐藤 「そういえば支店長がS C Mって製造と販売のアライアンスだといってたなあ。アライアンスということは互いの幸せを考えるということですよ」

山田 「もっといえば、うちと小売店が合併して1つの会社になったと考えれば良いんじゃないか」

佐藤 「そうなるとE O Qが使えるんだ。つまり最小費用を持つんで、ここが適正サイズになるんだ」

山田 「ただそのためには在庫費用と発注費用がきちんと計算されないとだめだらう。在庫費用はそれほど難しくないが、問題は配送費用だ。配送回数が変わると配送費用がどう変わるか、つまり1回あたりの配送費用をきちんと出す必要があるだらう。しかも配送は今考えているように単一商品でなく、いろいろな商品をまとめてやるだらう。だから計算が大変なんだ」

佐藤「そうか、それが今うちでSCMと一緒に考えているABC*か。だから1つ1つの仕事の費用を出さなくては出来ないわけか。合併と違ってアライアンスなんだから、売る方と買う方の互いが納得できる数字を出さなきゃだめなんだ。そしてそれを取引価格に反映させていくべきなんだ。よし、MCスーパーとはSCMとABCで最小費用の適正発注ロットを探すぞ」

 ABC

Activity Based Costingの略。活動基準原価計算と訳される。企業活動における各作業にかかった費用を計算していくこと。

佐藤は早速MCスーパーの松本バイヤーを訪ねた。

佐藤「発注ロットについてお話したいんですが」

とってEOQについて一通り説明した。

松本「主旨はわかった。前に私もEOQは勉強したことがあるけど、EOQは毎日毎日売れる量がはっきりしてる商品が対象なんじゃないの。もっといえば発注ロットは私と貴方のような担当者じゃなく、お宅の会社とうちの会社で基本的な取り決めをするもんじゃないの」

佐藤は心の中で「そうだよな。だから本社で委員会作ってたんだ。その委員会で言わなきゃいけないんだ」と思った。

欠品、品切れ

店舗などで消費者の需要に対して商品在庫がない状態。

松本「この店で今一番大切なことは欠品*をなくし、かつ売れ残りを出さないことだよ。そっちが先だよ。発注ロットの大小は会社と会社で話し合ってもらい、我々はこのロットを発注条件として考えれば良いんでしょう。11個欲しい時6個ロットなら2ロット買うということでしょう。私の加工食品部門で特に悩んでいるのはペットボトルの飲料なんだ。一体何個在庫すれば良いのか困っているんだ。毎日、毎日1つ1つの商品について考えていたら時間がいくらあっても足りないし、だからといってドンブリ勘定でというわけにもいかず困ってるんだ。私としては基本的なコンセプトというか、考え方を見つけないんだが、今1つわからないんだ。ペット飲料にはお宅の商品もあるんだから一緒に考えてくれない？」

佐藤「わかりました」

松本「これがうちが扱っているペット飲料の前期の日販量なんだけど、ちょっと見てくれるかな」 (図4-5)

佐藤はとらえずミドリ食品の主力ペット飲料である、ウォータースーパー350mlから考えてみることにした。

9月月報

8月月報

7月月報 (15.07.01~15.07.31) 中央店: ペット飲料

商品名	1日(火)	2日(水)	3日(木)...
ウォータースーパー 350ml	25	32	24
ウォータースーパー 500ml	18	12	16
川の純水 350ml	7	9	11
川の純水 500ml		
⋮	⋮	⋮	

図4-5

「とりあえず日販量の平均を出してみよう。表計算ソフトに入れてと……。平均日販は約27か、結構出てるなあ。そうだ前に勉強した標準偏差を出してみよう、約8か。で、どうすりゃいいんだ。ここはやっぱり数学の土屋さんしかいないな」

佐藤は土屋の所へ行って状況を説明した。

土屋「佐藤の次のテーマは在庫か。これがわかれば前にやったコンスタントのことももっとわかるようになるよ。まず在庫をやる前に正規分布について勉強しろ」

佐藤は『誰でもわかる在庫管理のやり方』を読んだが、正規分布という言葉は出て来ないので、また『誰でもわかる統計の本』を引っぱり出した。しかし正規分布の前の確率分布という所で止まってしまった。(図4-6)

確率分布

変数 x のとりうる値が $x_1, x_2, x_3 \dots$ と限られており、 x がこれをとる確率が $p_1, p_2, p_3 \dots$ と決まっている時、 x を離散確率変数という。サイコロの目などをイメージするとよい。

一方、連続確率分布とは、この変数 x のとりうる値が連続数で、その確率が決まるものをいう。小学校の身体検査で身長を測る時などをイメージするとよい。

連続確率分布において変数 x が $x_1 \sim x_2$ までの値をとる確率は積分して次のように求められる。

$$p(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} f(x) dx$$

この時 $f(x)$ を確率密度関数という。

図4-6

「最初のサイコロはわかるぞ。目が6通りあるので、目が1となる確率は、2となる確率はということだな。次は連続確率分布か、小学校の身体検査か、忘れてしまったなあ。身体検査で身長を測ることと確率の関係は？そうか、きっと日本中の小学生が母集団で、身体検査の身長が標本だ。連続数というのは、そうか身長には140.3cmってあってつながっているけど、サイコロの目には2.5がなくてつながっていないということか。小学生の身長を特定の小学校の身体検査で測って、日本全国の小学生の身長が140cmである確率を出そうということか。何となくわかるぞ」(図4-7)

「しかしこの積分はまいったぞ。もう土屋さんにたよるしかない」

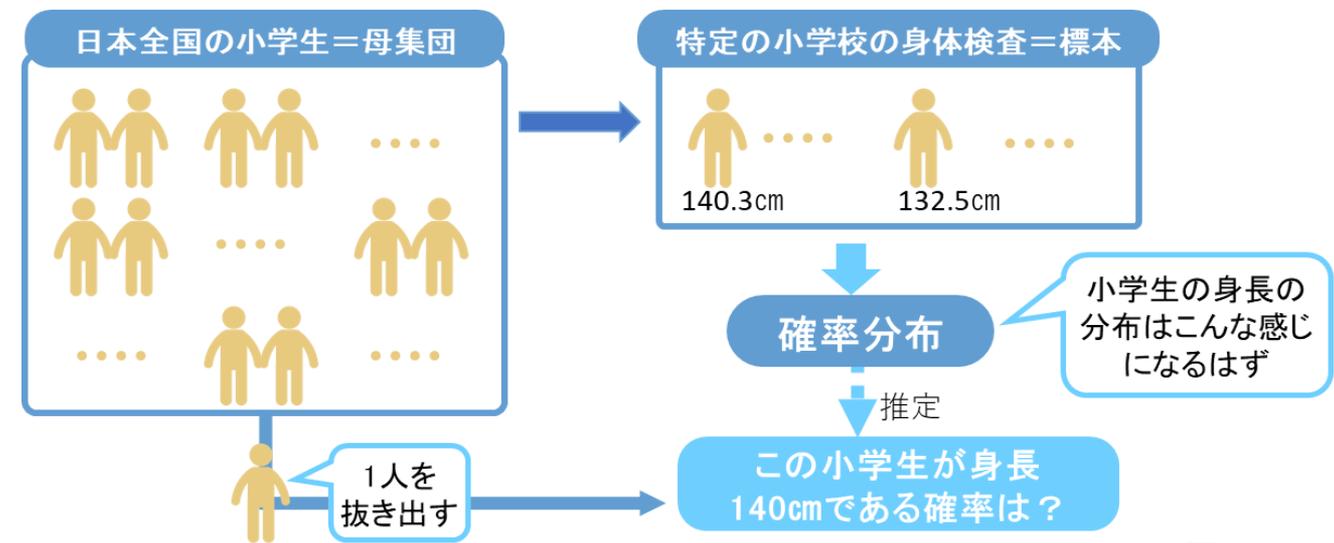


図4-7

翌日土屋にこの件を質問した。

土屋 「まあ、そこまでわかれば上出来だ。ところで佐藤はヒストグラムって知ってるか」

佐藤 「知りません」

土屋 「度数を表す帯グラフだ。例えば小学生の身長をランクに分けて考えるとす。そしてこのランクに入った人の数をカウントしていったものが度数だ。つまりそのランクに入る標本の数だ」(図4-8)

ランク	度数
120cm以下	6
120～130cm	11
130～140cm	17
140～150cm	32
150～160cm	19
160～170cm	10
170cm以上	5
合 計	100



図4-8

土屋 「これをグラフにしてみよう」(図4-9)

土屋 「これがヒストグラムだ。ところで160cm~170cmの人が100人のうち10人いるが、全体の何%だ？」

佐藤 「10%です」

土屋 「全国の小学生で160cm~170cmの間にある人は何%いると考えるべき？」

佐藤 「このデータからすると10%です」

土屋 「そう考えるべきだよな。じゃあ、全国の小学生から1人選んで身長が160cm~170cmである確率は？」

佐藤 「10%と考えるしかありません」

土屋 「ということはこのヒストグラムの高さと考えていいわけだ。棒の高さは確率だ。これをよく覚えておいてくれ」

土屋 「もっと多くの小学生に身体検査をやってランクを細かくするとどうなる？」

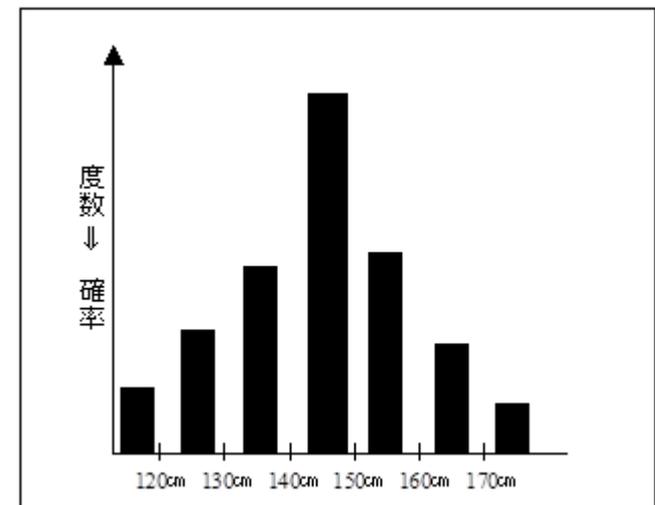


図4-9

佐藤 「こんな感じですか」 (図4-10)

土屋 「身体検査をもっともっと多くの全国の小学生にやって、データのランクをもっともっと細かくしていくとどんな感じになる」

佐藤 「だんだん棒がつながっていきますよね」 (図4-11)

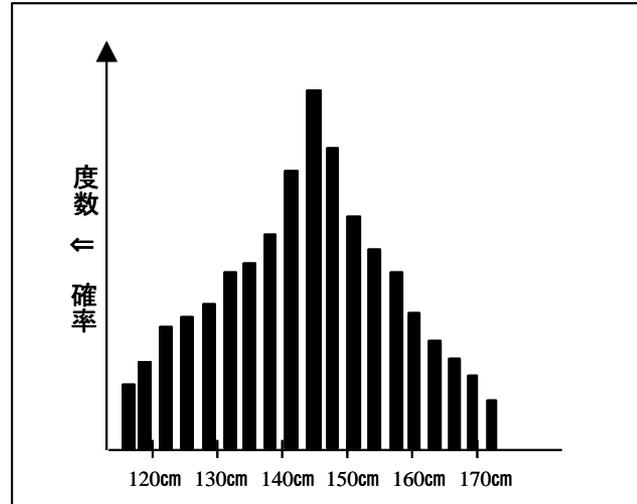


図4-10

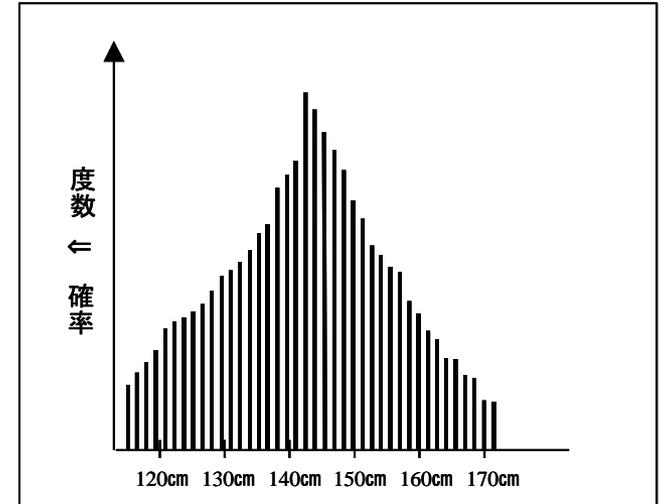


図4-11

土屋 「この棒の頭をつないでみるとどうなる？」

佐藤 「こんな感じですか」 (図4-12)

土屋 「ところで横軸はなんだ？」

佐藤 「小学生の身長です」

土屋 「たて軸は？」

佐藤 「確率です」

土屋 「これが確率密度関数というグラフだ。この細かい棒が確率の密度だ。全国の小学生から1人選んで身長が160cm~170cmになる確率はどうなる？」

佐藤 「わかりません」

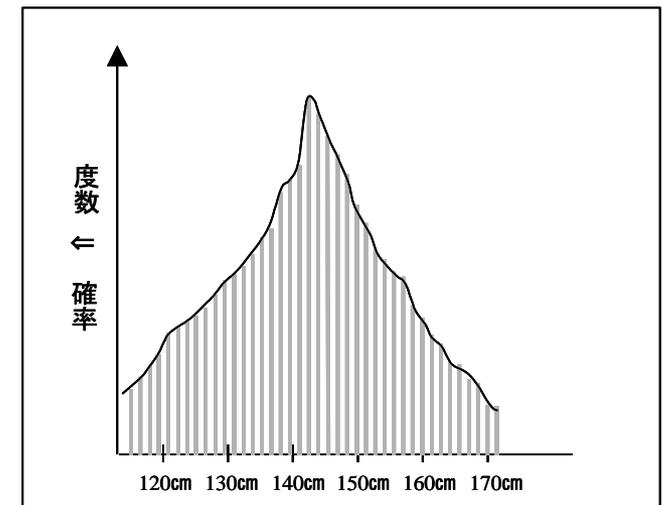


図4-12

土屋 「さっきやったろう、160と170のところ
たてに線を引いてみる」 (図4-13・上図)

土屋 「もう1度一番最初のヒストグラムを見て
みる」

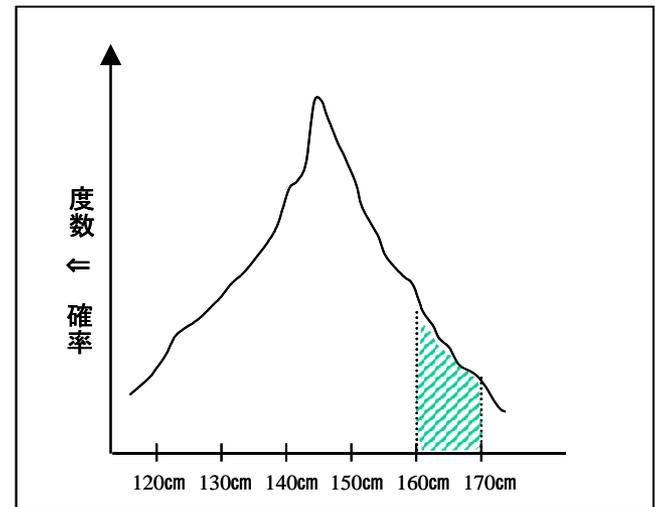
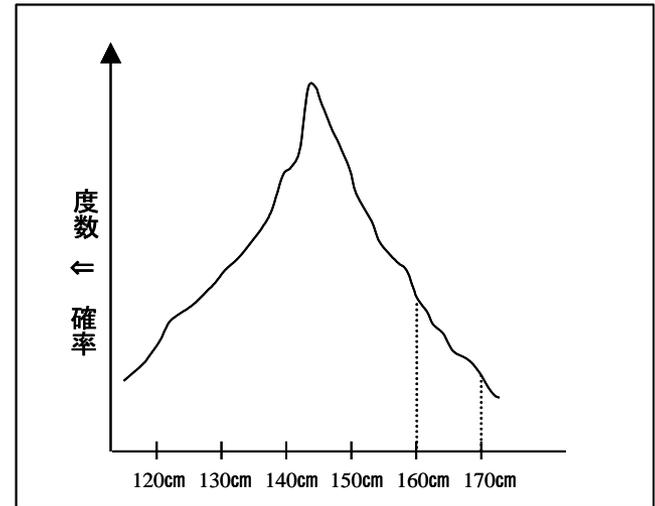
佐藤 「わかりました。高さが確率ですね」

土屋 「160cmの所と170cmの所では高さが違うだ
ろう」

佐藤 「平均しましょうか」

土屋 「160cm～170cmの間にたて線が無数にある
ぞ。そんなの平均とれないぞ。ヒストグラム
では確率は高さというよりも棒の面積と
考えても良いんじゃないか」

佐藤 「わかった。この部分の面積でしょう」
(図4-14・下図)



土屋 「正解だ。それが君の読んだ本の式だ。つまり積分だ。積分は面積を出すことなんだ」

佐藤 「でも面積を計算するのは大変ですね」

土屋 「心配するな。コンピュータが全部やってくれる」

佐藤 「そうするとこれって便利ですね。何cmから何cmでも、何cm以上でも、何でも確率が出ますよね。ところで身長が140cmちょうどになる確率はどうなるんですか？」

土屋 「ジャスト140cmなんてことはないだろう。例えば小数点第1位を四捨五入して140cmなら、139.5cmから140.4cmだ」

佐藤 「そう考えればいいのか」

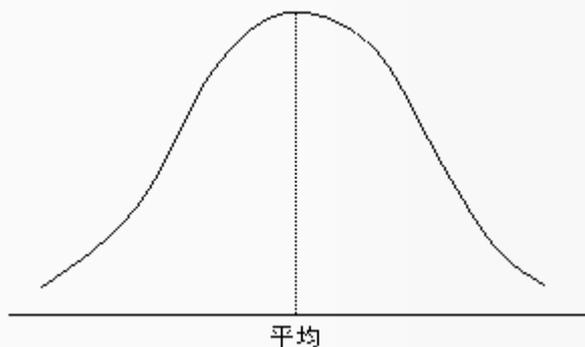
土屋 「これが確率分布だ。次は正規分布だ。もう一度勉強して来い」

佐藤は家で「誰でもわかる統計の本」の正規分析の所を読んだ。

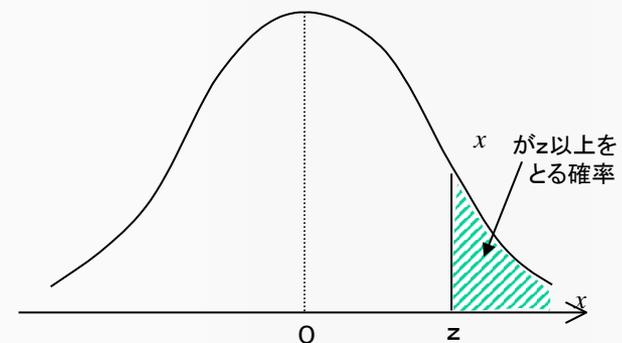
図4-15

正規分布

正規分布とは確率密度関数が下図のような平均を中心として、左右対称のつり鐘型のグラフを示すものである。数学的に非常に扱いやすく、さまざまなデータ分析に用いられる。



平均0、標準偏差1の正規分布においては次のような確率が表で与えられる



「平均が真中で一番高いのか。何か小学生の身長も真中あたりが高くてこんな感じだったなあ。下の z 以上というのは斜線の所の面積だな」

翌日佐藤は土屋に質問した。

佐藤「正規分布は数学的に非常に扱いやすいってどうゆうことなんですか」

土屋「それには3つの意味があるんだ。1つは平均を中心にして左右対称ということだ」(図4-17)

土屋「つまりその値が平均より大きい確率と小さい確率が共に $1/2$ だ。これは非常に直感的だ。小学生の平均身長が 140cm なら、 140cm より低い小学生と高い小学生が同数になるということだ」

佐藤「ありがちで、わかりやすいですね」

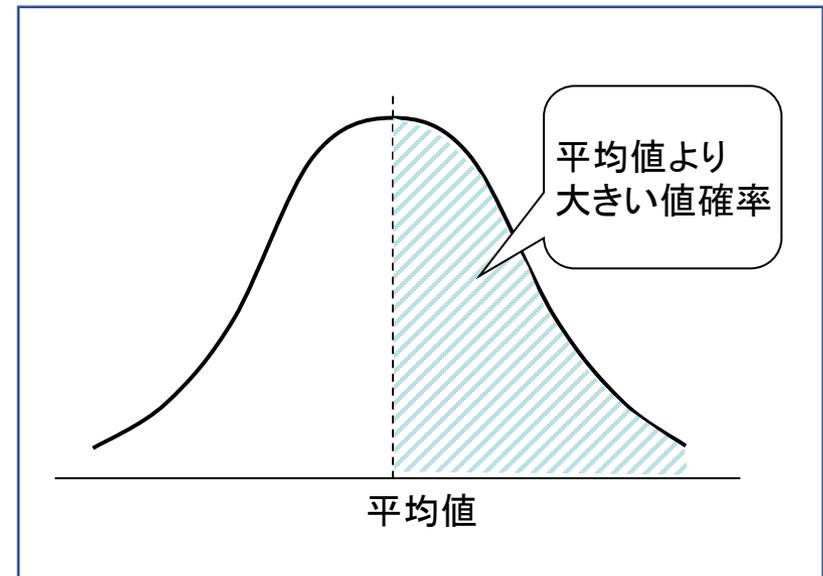


図4-17

土屋「2つ目は表を使って面積、つまり確率がすぐわかることだ。例えば小学生の平均身長が140cmで、170cm以上の人の確率を知りたい時はこの部分の面積を表で計算できるんだ」(図4-16, 4-18)

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4206	.4220	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4358	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4601	.4610	.4619	.4628	.4637
1.8	.4646	.4655	.4664	.4673	.4682	.4691	.4700	.4709	.4718	.4727
1.9	.4736	.4745	.4754	.4763	.4772	.4781	.4790	.4799	.4808	.4817
2.0	.4826	.4834	.4843	.4851	.4859	.4867	.4876	.4884	.4892	.4900
2.1	.4908	.4916	.4924	.4932	.4940	.4948	.4956	.4964	.4972	.4980
2.2	.4988	.4996	.5004	.5012	.5020	.5028	.5036	.5044	.5052	.5060
2.3	.5068	.5076	.5084	.5092	.5100	.5108	.5116	.5124	.5132	.5140
2.4	.5148	.5156	.5164	.5172	.5180	.5188	.5196	.5204	.5212	.5220
2.5	.5228	.5236	.5244	.5252	.5260	.5268	.5276	.5284	.5292	.5300
2.6	.5308	.5316	.5324	.5332	.5340	.5348	.5356	.5364	.5372	.5380
2.7	.5388	.5396	.5404	.5412	.5420	.5428	.5436	.5444	.5452	.5460
2.8	.5468	.5476	.5484	.5492	.5500	.5508	.5516	.5524	.5532	.5540
2.9	.5548	.5556	.5564	.5572	.5580	.5588	.5596	.5604	.5612	.5620
3.0	.5628	.5636	.5644	.5652	.5660	.5668	.5676	.5684	.5692	.5700

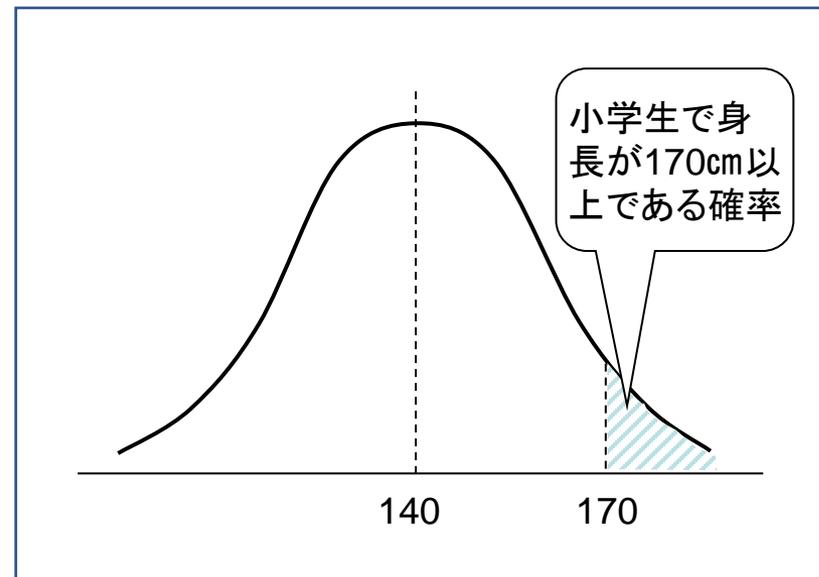


図4-16

図4-18

佐藤「あの一、例えば身長130~160cmの確率を知りたい時はどうするんですか」

土屋「その時は線の下をまず①~③の部分に分ける」

と書いて紙に書いて説明していった。(図4-19)

土屋「求めたいのは②の面積だな。ところで①、②、③の面積を足すといくつだ」

佐藤「①から③を足すということは、すべての確率を足すわけですよ。確率は全部足すと1ですよ」

土屋「③はさっきの表でわかるよな。①は？このグラフは左右対称だぞ」

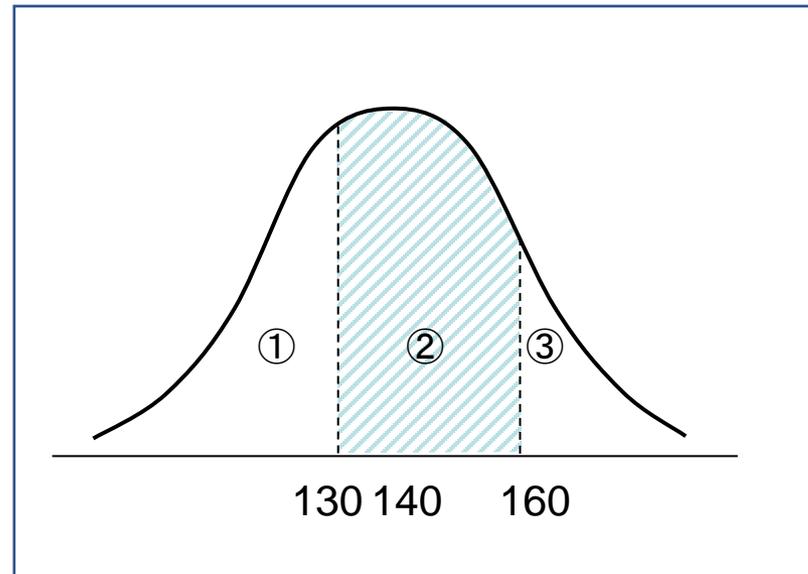


図4-19

佐藤「そうか、身長から右側に10の所、つまり150の所の表を見ればいいんですよね」 (図4-20)

土屋「1から①と③を引けば②が残る。つまり130cmから160cmをとる確率だ」

佐藤「便利ですね」

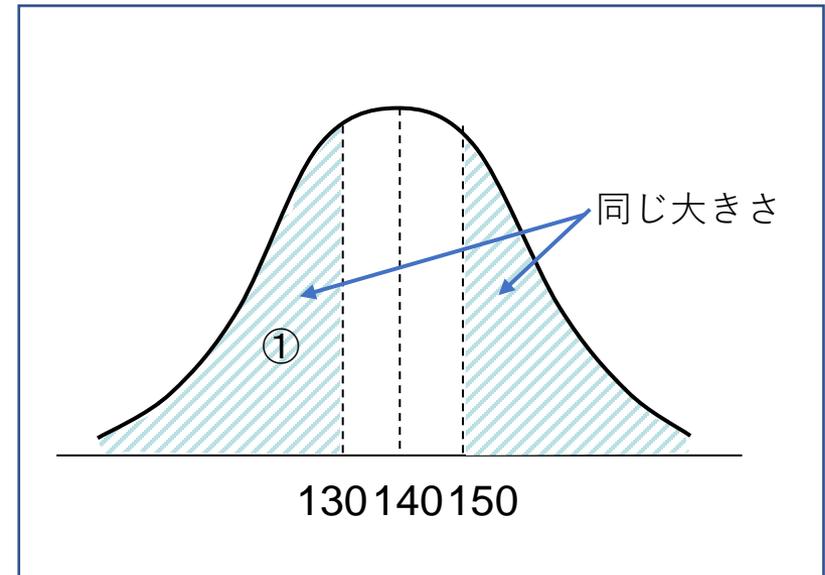


図4-20

土屋 「3つ目の使いやすさは、世の中の現象で結構つり鐘型の感じのものが多いことだ。小学生の身長なんかもこんなつり鐘型と考えても変じゃないよな」

佐藤 「はい。前のヒストグラムとよく似ていると思いました」

土屋 「そして今企業がもっとも注目し、佐藤がずっと悩んでいる需要も正規分布と考えられるんだ」

佐藤 「本当ですか？誰か証明したんですか？」

土屋 「証明なんてできるわけないだろう。需要なんて人間が買いたいと思う気持ちだぞ。科学的に証明できるわけないじゃないか。ただ需要の結果といえる販売個数を、小学生の身長のようにしてヒストグラムを作ると大体つり鐘型になるんだ」

佐藤 「経験則ってやつですね。でもちょっと待って下さい。需要を販売個数で考えると連続量ではないですよ。身長には130.34cmとかありますけど、個数にはないですよ」

土屋 「このグラフを遠くから見てみる」 ([図4-21](#))

土屋 「1つ1つの点が個数だが、つながっているように見えないか」

佐藤 「見えるような、見えないような」

土屋 「見えれば便利だぞ。さっきの表が使えるぞ」

佐藤 「何だか見えてくるような気がしてきました」

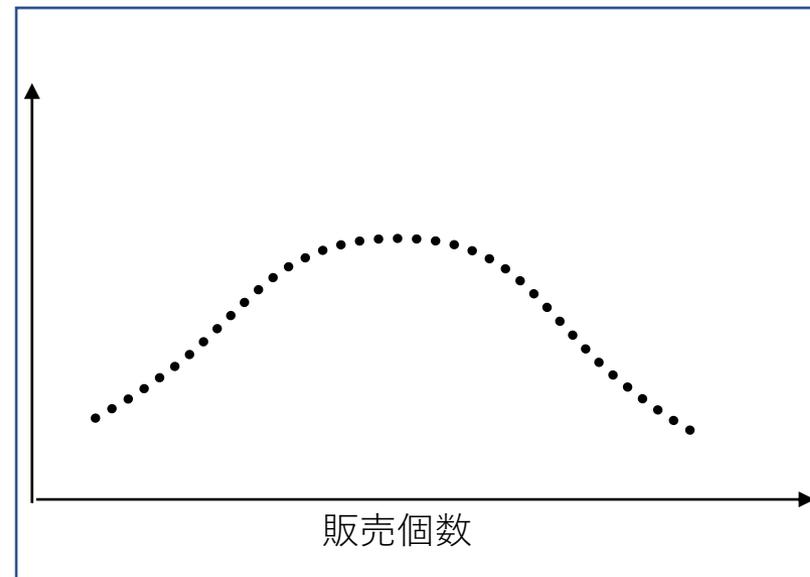


図4-21

土屋 「いよいよ本番だ」

佐藤 「まずはウォータースーパー350mlの店舗での在庫を考えたいんです。在庫とは店に売れる量だけ置いておけば良いんですよね。まさに売れる量は需要であり、需要予測ですよね。需要は正規分布と考えることにすると。あっ、標本の統計量もちゃんと計算しました。1日平均27、標準偏差が8です。この時はどの表を使えば良いんですか？」

土屋 「正規分布表は1つだ。平均0、標準偏差1しかない」

佐藤 「でもどうやって使うんですか」

土屋 「何てことはないんだ。販売個数のすべてのデータから27を引くと平均はどうなる？」

佐藤 「平均は27小さくなりますから0です」

土屋 「すべてのデータを8で割ると標準偏差はどうなる？」

佐藤 「もしかして1/8になるんですか？」

土屋 「そうだ。これで標準偏差は1だ。ウォータースーパー350mlの販売個数から27を引いて8で割ったデータは、平均1で標準偏差0だ。そして後で27を足して8をかけりゃいいんだ」

佐藤 「かしこい！」

土屋 「ところでその在庫量の予測対象期間は？」

佐藤 「何ですかそれは？」

土屋 「じゃあ聞き方を変えよう。どうゆう風にMCスーパーは発注してるんだ」

佐藤 「その発注のやり方自体を考え直したいんです」

土屋 「わかった。じゃあMCスーパーにとって最善の発注方法はどんな感じだ」

佐藤 「それは前日の閉店後、売れる量を考えて注文し、注文されたものが当日の朝、開店前に納品される感じでしょう」

土屋 「じゃ、まずそれでいこう。MCスーパーは何日分在庫を持てば良いんだ」

佐藤 「1日分ですよね」

土屋 「じゃあ、売れ残りが無いとして、何個発注すれば良い？つまり毎朝在庫が何個あれば良い？」

佐藤 「平均27個ですから、27個でしょう」

土屋 「よく考えろ。グラフを書いてみろ」 ([図4-22](#))

土屋 「日販量が27個より大きい確率は」

佐藤 「1/2です」

土屋 「これで良いのか」

佐藤 「よく考えたら2日に1回欠品してしまいます。いくつにしましょう？MCスーパーに言わせれば欠品なんて問題外だというのでしょうか。ということはこのグラフの『右はし』が下の線とくっついた所にすれば絶対欠品しませんよね。でも何だか随分在庫が多いような気がします」

土屋 「残念ながらこのグラフに『右はし』というものがなく、永遠に下の線とはくっつかないんだ」

佐藤 「ということは、欠品を出さない在庫と
いうのはないんですか」

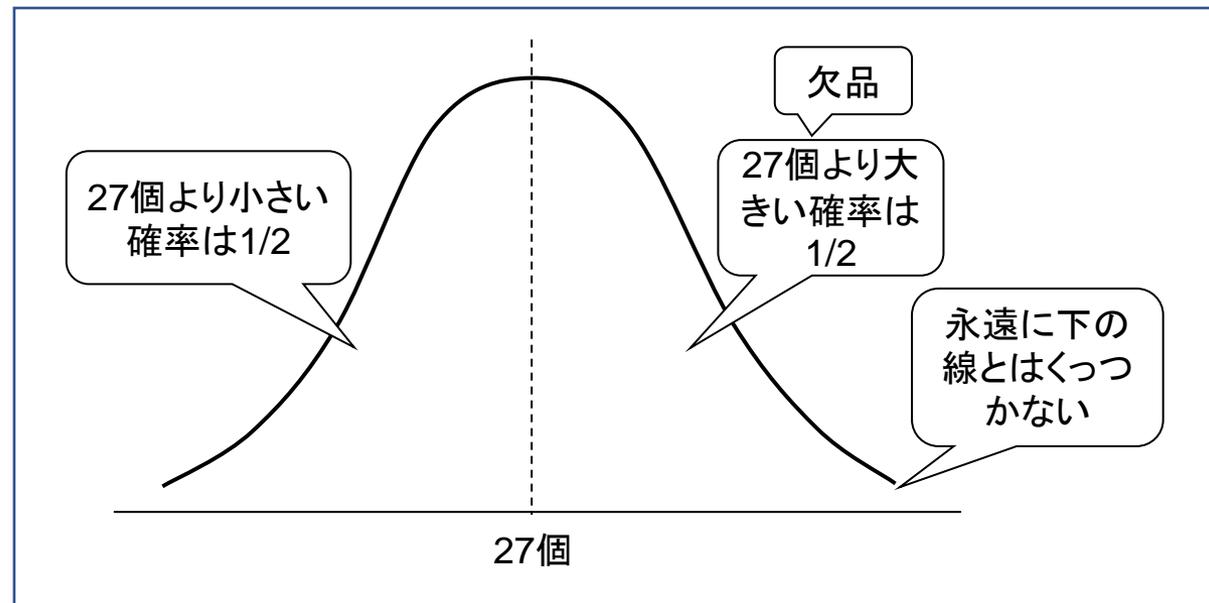


図4-22

土屋 「数学的に考えればそういうことだ。例えば1日平均2個しか売れてなくても、何かの拍子に100個売れる可能性は0とはいえないだろう。だからMCスーパーに『欠品を起こさない』ようにはできないということを理解してもらうんだ。これは欠品を起こしてから言えば言い訳だ。でも発注する前に在庫を一緒に考えているなら納得してくれるはずだ。MCスーパーにとっても、めったにない100個売れる時のために、在庫を抱えておくのはいやだろう。まあ、うちにとっては悪くない話かもしれないけど」

佐藤 「わかりました。結局いくつにしたらいいんですか」

土屋 「それが佐藤の悪い所だ。いくつにするかという答えよりも、どうやって適正在庫を考えるのかが大切だろう。そして何度もいっているように、佐藤が悩んできたことなんか昔から多くの人が悩んできたと思うことだ。そしてその中の賢い人が出した結論を使うんだ」

佐藤 「何だ、あるんだったら最初に教えて下さいよ。勉強しなくてすんだのに」

土屋 「ある。だけど考え方を理解しないで、納得しないで、数字を出したって、お客様も理解してくれないし、納得してくれないぞ。これは一言でいえば安全在庫だ。ここからは在庫管理の本に書いてあるはずだ」

佐藤は今度は「誰でもわかる在庫管理のやり方」の本を広げた。(図4-23)

需要変動に対してバッファとなる在庫を安全在庫という。適正在庫は次のような式で求められる。

$$\text{適正在庫} = \text{需要平均値} + \text{安全在庫}$$

ここで、「安全在庫 = $k \times$ 標準偏差」となる。

k は安全係数とよばれ、許容品切れ確率によって決まる。許容品切れ確率10%の時、安全係数は1.28となる。

図4-23

「許容品切れ確率10%ということは、10回に1回に欠品を許すということかな。なんだか多いなあ。でもこれしかないからこれでやってみよう。需要平均値は27でOKだ。標準偏差は8だから安全在庫は1.28をかけて10個か、ということは37個が適正在庫か」

佐藤「土屋さん、適正在庫が出ました。でもよく意味がわかりません」
土屋はまたグラフを書いて説明した。(図4-24)

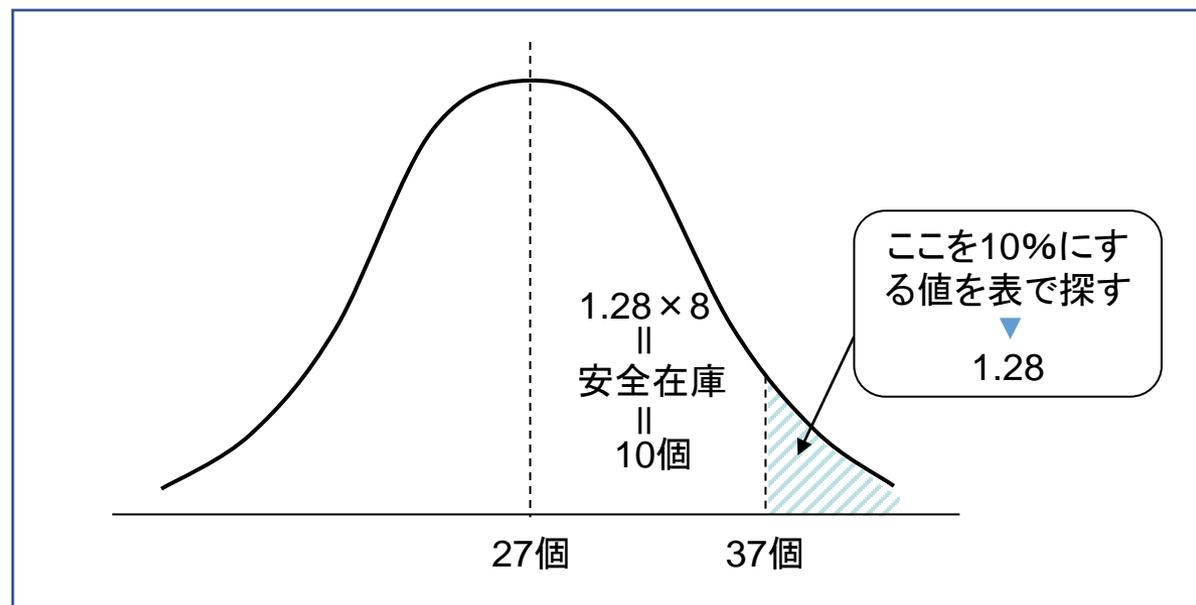


図4-24

土屋 「許容品切れ確率とは、この右はしの斜線の部分の面積を10%、つまり0.1になるようにすることだ。さっきの逆で面積が0.1になるような数字を表から探すことなんだ。この値が1.28だ。佐藤が持っているの本のzにあたる所だ ([204ページ図4-15](#))。これは標準偏差1の時だから、8の時は8倍するんだ。これで10になる。つまり右へ10個増やさなくてはならないということだ」

佐藤 「でも10%ってことは10回に1回品切れを起こすということでしょう」

土屋 「数学的にはね。でも正確に正規分布どおり消費者は買物するわけじゃないだろう。正規分布というのはもっともデタラメに物事が起こると考えているんだ」

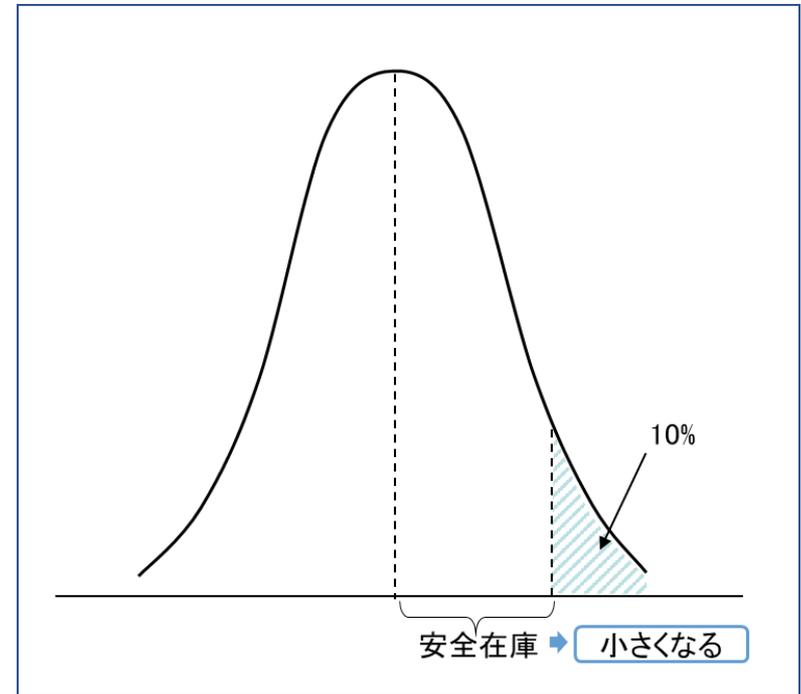
佐藤 「デタラメというのは？」

土屋「まあ、消費者が昨日何を何個買ったか、他の人は何を買ったか、ライバル店がどうなっているかといったことを一切考慮しないで買うようなものだ。ところが実際はそんなにデタラメではなく、ある程度の規則性はあるだろう。例えば毎日1個ずつ買いつづける人だっているだろう。完全にデタラメだと考えると、現実の世界よりはどうしても安全在庫は大きく出てしまうんだ。いってみれば正規分布で許容品切れ確率が10%という『欠品はあまり起こらない』位かな、もちろん5%にしてもOKだ。この時は『めったに起こらない』位かな。1%だと『ほとんど起こらない』位だ」

佐藤 「今気づいたんですけど、安全在庫って標準偏差で決まりますよね。ということは標準偏差が小さくなると安全在庫は減りますよね」

土屋 「標準偏差が小さいというのは、こんなグラフになるんだ」

(図4-25・左図)



佐藤 「そうか標準偏差が小さい、つまりコンスタントに売れるというのは、在庫を削減する効果があるんだ。前に考えたうちのスーパーグリーンは標準偏差が小さいんで、少ない在庫で済むんだ。わかってきたぞ。コンビニはコンスタントに売れるものしか置かず、アイテムあたりの在庫を落として、アイテム数を増やしたんだ」

土屋 「そうか、俺もそこまでは気づかなかった。佐藤、それならROIの考え方を使え」

佐藤 「ROIって何ですか？」

土屋 「リターン・オン・インベストメントだ」

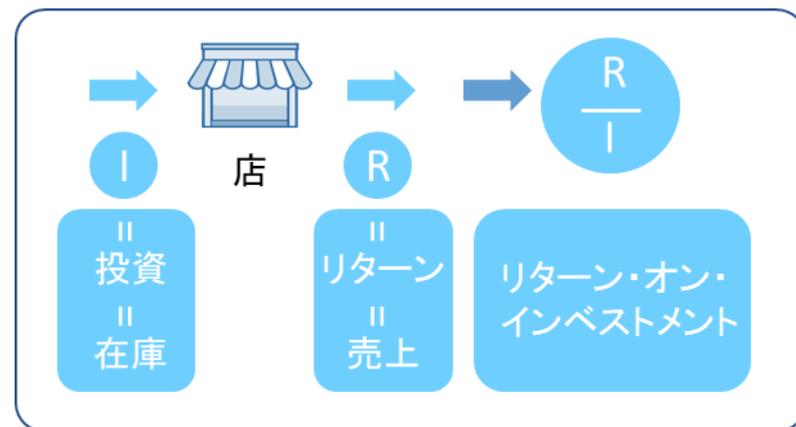
とって紙に絵を書いた。

図4-26)

土屋 「いいか、インベストメントつまり投資、これがIだ。

投資してリターンRを得るんだ。この時その投資の妥当性をROIで見るんだ。Iの上にRが乗っているだろう。これがオンだ。中学の英語でやったろう。屋根の上の鳥はバード・オン・ザ・ルーフだ。ビジネスの収益性、生産性はすべてこれだ、覚えておくと便利だぞ」

佐藤 「この場合は投資が在庫で、リターンが売上かな。つまり売上が在庫でわるのか。あれ、この式前に見たことがあります。在庫を金額で考えれば商品回転率だ。時々販売レポートで見かけるけど意味がわからなかったんです。商品回転率とは在庫投資の収益性か。つまりコンスタントに売れる商品はIの在庫を小さくして、商品回転率を上げるんだ。コンビニは商品回転率を上げ、大成功したと聞いたことがあるぞ」



土屋「トヨタもそうだ。工場にいた時よくいわれたなあ。工場では在庫回転率というんだ。、『在庫回転率を上げるというのは居酒屋で客の回転率を上げるのと同じだ。回転率が良いと活気が良くなり、現金のまわりも良くなる』と何度も何度も聞かされたなあ」

佐藤「そうですね、商品回転率を上げると、店では現金になるスピードが上がり、店員も活気づきますよね。居酒屋で客の回転が良いと、つまみも新鮮だし、店員もイキイキしてますよね」

佐藤 「ところでMCスーパーが2日分の在庫を持つようにしたらどうなるんですか。単純に在庫は1日の時の2倍になるんですか」

土屋 「需要の平均は2倍になるけど、標準偏差は $\sqrt{2}$ 倍なんだ。約1.4倍かな」

佐藤 「そうすると平均が54、安全在庫が14個で68個か。1日分が37個だから1.8倍くらいか」

佐藤はMCスーパーでプレゼンテーションをしていた。

「いただきましたレポートから各アイテムの日販数の平均、標準偏差をもとに1日分の適正在庫を出しました。ただし許容品切れ確率は10%で計算しています。10%という線は直感的にはあまり欠品が起こらない状態といえます。

この確率を下げていくと在庫は増えていきます。この条件でウォータースーパー350mlの適正在庫は37個となります。1ロットが6個ですので、手持ち在庫を引いてロット数を出すこととなります。また2日に1回の発注として、2日分の在庫を持つとすると68個となり、安全在庫の関係で2倍より小さくなります」

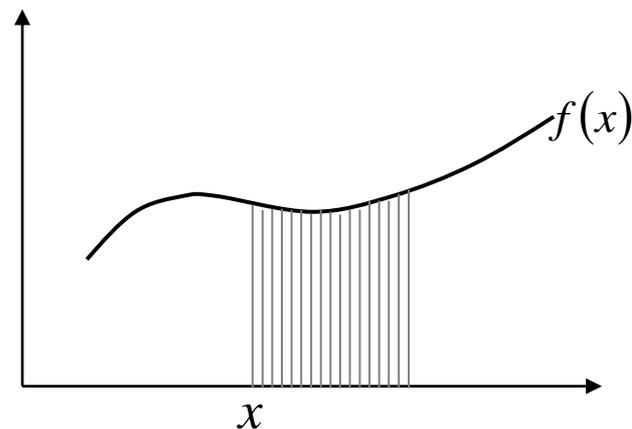
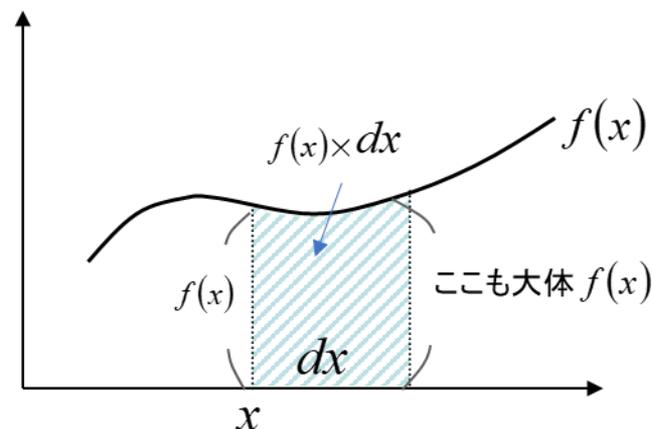
「一方、毎日発注すると2日に1回発注するのでは1個あたりの配送コストが異なりますので、納入単価は毎日の方がどうしても高くなります。前にお話したE O Qの考え方で貴社および当社にとっての適正発注サイクル、それに伴う適正在庫を検討させていただければと思います。

ちなみに、標準偏差の低下は安全在庫を落とす効果があります。そういう目で、各商品アイテムについて、平均日販とともに標準偏差も見て、店舗でのお取扱いの有無も検討して下さい」

積分

本文でも述べたとおり、積分とは図形の面積を出すために考えられたものです。積分では f という不思議な文字を使います。 f はsum(和)という単語の頭文字sから来たものです。下のグラフのような $f(x)$ という曲線が作る斜線の部分の面積を出すことをいいます。

上のグラフで x からちょっとだけ右へいった(これを dx と表します。デルタの d です) 所までの面積を出したいとします。面積は「よこ×たて」です。よこは dx です。たては右の線と左の線で少し違います。左の線は $f(x)$ です。右の線も dx が小さいので大体 $f(x)$ です。ということは長方形となり、面積は(算数では \times をよく省略するので) $f(x)dx$ です。 dx がかなり小さいと x の近くの細かい線のようなものになります。



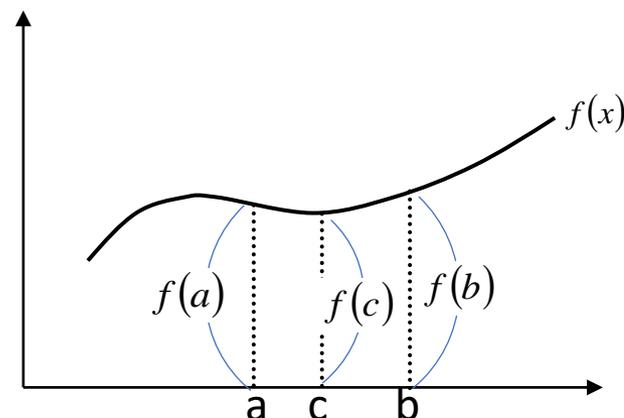
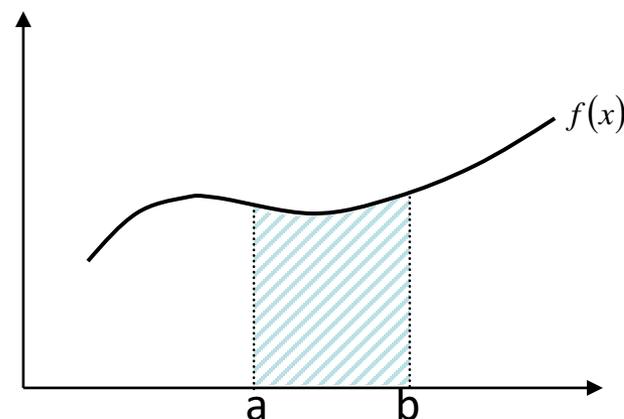
そしてこの線を足していくと次の図のaからbまでの面積になっていきます。

この「 $f(x)dx$ という細かい線をaからbの部分まで足して、面積を出すこと」を「 $f(x)$ をaからbまで積分する」といい、「 $\int_a^b f(x)dx$ 」と表します。 f は和です。積分とは分(細かい線)を積み上げる(かけ算する)という意味です。

積分でなぜ確率が計算できるかを考えてみましょう。分(細かい線)を積むという感覚をわかって下さい。

$f(x)$ が確率を表す曲線と考えると、 a をとる確率は $f(a)$ 、 b をとる確率は $f(b)$ となります。さらにその間にある c では $f(c)$ です。この線をaからbまで全部足せば、 x がaからbまでの値をとる確率となります。

つまりaからbまでの値をとる確率は $\int_a^b f(x)dx$ です。



第 5 章

数字で説得する

ミドリ食品はトマトにかけるための新しい調味食品である「トマトソルト」を開発した。テレビコマーシャルを集中的に打っているのだが、今1つ売上が伸びず、本社マーケティング部では悩んでいた。マーケティング部員たちは、手分けして小売店舗でのトマトソルトの陳列状況を調査した。

「トマトソルトはほとんどの店舗で、ドレッシング売場においてあり、他のドレッシングよりパッケージ*が小さいので目立たない。何か埋もれてしまっている感じだ。主婦の認知度*が低いのではないか」

「トマトソルトなんだから青果売場のトマトのわきに置いてくれるほうがいいんだけど」

「トマトソルトをわきに置けば、トマトの売上も伸びるんじゃないか。子供たちの野菜不足が問題になっているし、トマトソルトをかければトマトの青臭さが消えるし、簡単だし夕食のサブメニューとしては悪くないはずだ。しかもサブメニューだから毎日食卓に少しずつ出してもおかしくないしね。」

「よし、小売店舗へトマトとのクロスマーチャンドライジング*の提案をして反応を見よう。全支店長に依頼のメールを打とう」

S支店の中村支店長はマーケティング部から届いたメールを見ていた。

中村「クロスマーチャンドライジングか。そりゃうちとしては売れないで悩んでいる『トマトソルト』をトマトのわきに置いてくれりゃうれしいよ。でもふつう生鮮のバイヤーがいやがるよなあ。トマトの売上が伸びるとも限らないし、大体青果のわきに、はみ出し陳列*なんてなかなかやらせてくれないよなあ。まあ明らかにトマトの売上に効果があることがわかっていれば別なんだが…。ということはどっかでやらせてもらうしかないか。そうだ、佐藤、ちょっと聞いてくれ。今いろいろなことを一緒に取り組んでいるMCスーパーにトマトソルトのクロスマーチャンドライジングを、無理言って頼めないかなあ」

佐藤「それならちょうどいいかもしれません。P店の加工食品部門にいた東さんが中央店の生鮮のバイヤーになったんですよ。P店の時に親しかったので頼んでみます」



パッケージ

商品包装のこと。商品が入っている箱、袋などをさす。

認知度

店舗などで顧客が商品があることを認知する度合い。

クロスマーチャンドライジング

商品部門を超えて考えることをクロスマーチャンドライジングと言う。一般に店舗での商品陳列は加工食品、生鮮、日用雑貨などの部門ごとに行われる。この部門を超えて陳列を行うことをクロスマーチャンドライジングと言うことが多い。この場合、関連陳列とも言う。トマト（生鮮）のわきにトマトソルト（加工食品）を置くことなど。

はみ出し陳列

定番の商品売場の脇などの通路に、はみ出して陳列すること。

佐藤はMCスーパー中央店青果売場にいた。青果売場は整然としており、加工食品の陳列は何もなかった。

佐藤「東さん、どうですか生鮮は？」

東「何かぱっとしないね。どう考えてもうちの店の利益源だし、金のなる木のはずなんだけどね。特に野菜が今1つだよ。近くの青果専門店が産地直送と称して、無農薬系を扱い始めて大分客を取られてしまったんだ。うちも無農薬の売場作ったんだけど、はかばかしくないんだ」

佐藤「東さん、うちでトマトソルトっていう新しい調味料を出したんですけど、トマト売り場の脇に置いてくれませんか？」

東「わきに置くといってもねえ」

佐藤「什器*の用意、セッティング、陳列まで全部うちがやります。加工食品の方にも了解はとってるんですけど」

東「そりゃ加食はいいよな。こっちにフェースが増えるし。でも生鮮側のメリットがないんじゃない？」

佐藤 「私はこれでトマトの売上が伸びると考えているんですよ。主婦に新しいトマトの食べ方を提案すれば、食卓にもう1品乗るような気がするんです。レタスかトマトかの選択購買※でなく、純粹な客単価アップにつながると思うんですけど。」

東 「まあ俺も来たばかりだし、指くわえて見ているのも何なんだからやってみるか。でもどうせやるなら成功させたいよね。来月の頭にうちで秋の大フェアやるから、その時やるか。」

佐藤 「そうですね。トマトの売価どうします？」

東 「まあ実験みたいなものだから、利益は度外視して思い切って下げるか。」

佐藤 「でも何か心配ですね。売り場でマネキン※立てて試食やったらだめですか」

東 「やろう。やれることは全部やらなきゃ意味ないよな」

 仕器

仕器とは日常使う家具のことをさすが、小売業では店舗で必要な棚、ワゴンなどの器具をさす。

選択購買

店舗において代替品と比較しながら購買すること。

マネキン

婦人服の陳列に使う人形のことともマネキンと言うが、ここでは小売店などに特定のメーカー、卸から派遣された販売員のことを言う。派遣店員と同意。

翌月、MCスーパー中央店では秋の大フェアの一貫として、トマトソルトの 프로모ーションも1週間に渡って実施された。POSデータからトマト、トマトソルトの売上実績を比較してみると次のようになった。(図5-1・[図5-2](#))

単位：円

	週	月	火	水	木	金	土	日	合計	前週比
トマト	10/6~12	8,260	7,420	7,750	8,450	8,960	12,460	15,380	68,680	1.56
	前週	6,380	4,460	6,800	4,960	3,720	8,860	8,720	43,900	—
トマト	10/6~12	2,000	2,750	3,280	2,250	3,500	6,250	5,750	25,780	6.78
ソルト	前週	600	900	0	300	300	200	1,500	3,800	—

図5-1

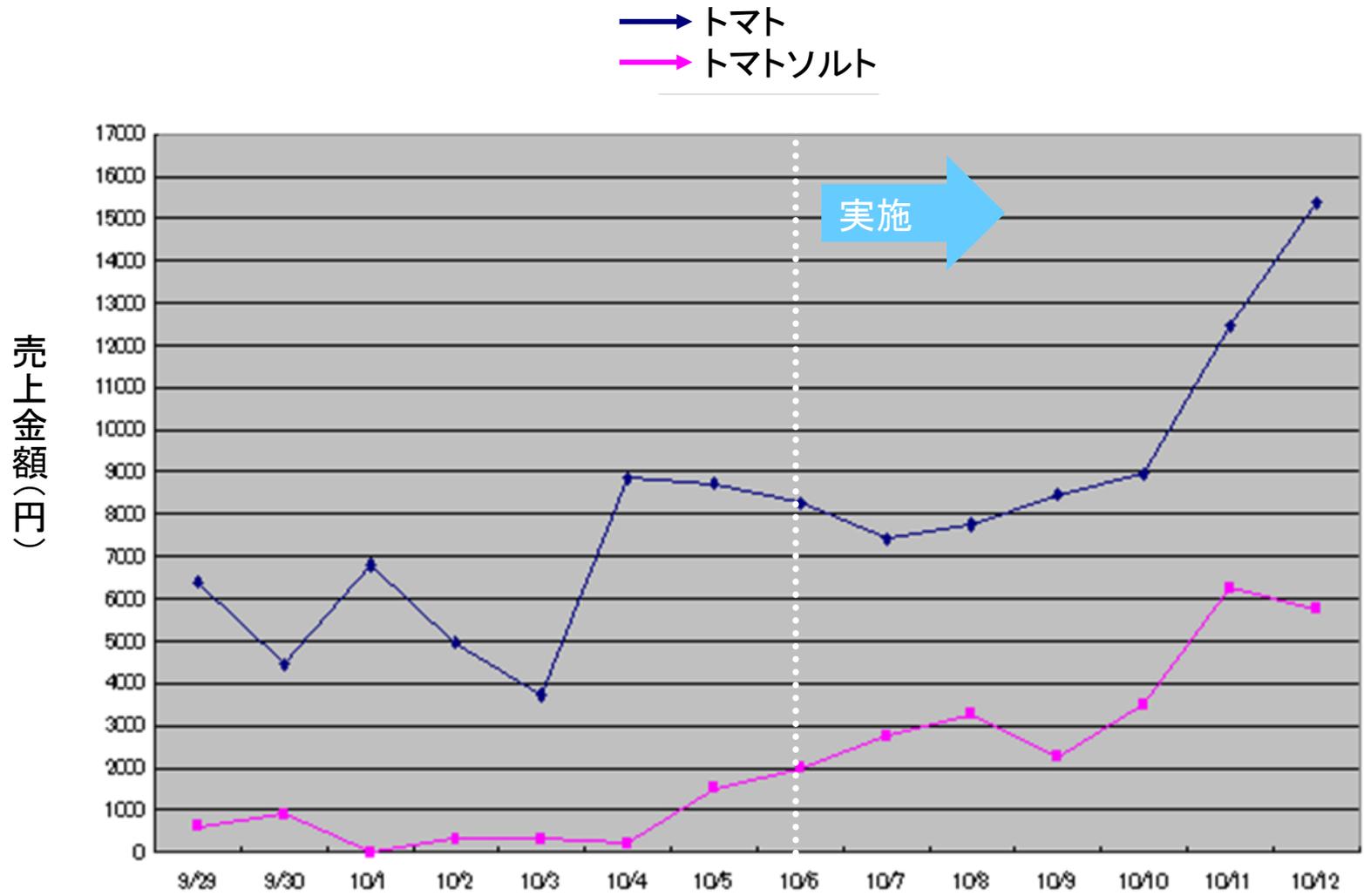


図5-2

佐藤 「出ましたね。トマトソルトは7倍近くですし、トマトだって1.5倍以上出ますよ。誰がどう見たって成果出てますよね」

東 「うまくいったよなあ。次はなにをやるうか？」

佐藤は中村支店長へもこの結果を報告した。

佐藤 「うまくいきました。MCスーパーの東バイヤーもすごく喜んでくれました」

中村 「そうかよかったなあ。トマトが1.5倍か。良い数字だなあ。まずはMCスーパーの他店に同じように提案していけ。それでうまくいったら本社で来月やる『セールス成功事例会』で報告して全国の店舗に展開させよう。これなら佐藤の社長賞だって夢じゃないかもしれないぞ」

佐藤はMCスーパーU店の生鮮バイヤーである青木氏の顔を浮かべていた。「青木さんは結構論理的だから、逆にこれだけの成功データがあれば乗ってくれるだろう」

佐藤はU店の青木を訪ねた。青木はパソコンでPOSデータを見ながら考えごとをしていた。

佐藤「青木さん、実は中央店でトマトとトマトソルトのクロスマーチャダイングやったら、驚くほどうまくいきました。この店でもやりましょうよ」

青木「そうやっていろいろなメーカーがいろいろな陳列提案をしてくるんで、やってみるんだが、やったその時はうまくいったような気がするけど、結局長続きしないんだよな。やる時肩に力を入れすぎて、やりすぎるんだよ。がんばりすぎて良い結果を出しすぎるんだ。良い結果がその提案のためだと思えば、結構そうでないことが多いんだよな。まあいいや、じゃその結果を見せて。このトマトの数字は日あたりの販売金額か。すごい額だな。待てよ。この時は確か秋のフェアやってなかった？」

佐藤「やってみました」

青木「それなら金額じゃ意味ないじゃない。来店客数が全然ちがうだろ。売上PI値*にして持ってきてよ」

佐藤は以前に受けたセールス研修のテキストを引っ張り出してきた。

「確か来店客数で割るんだっけ」 ([図5-3](#))

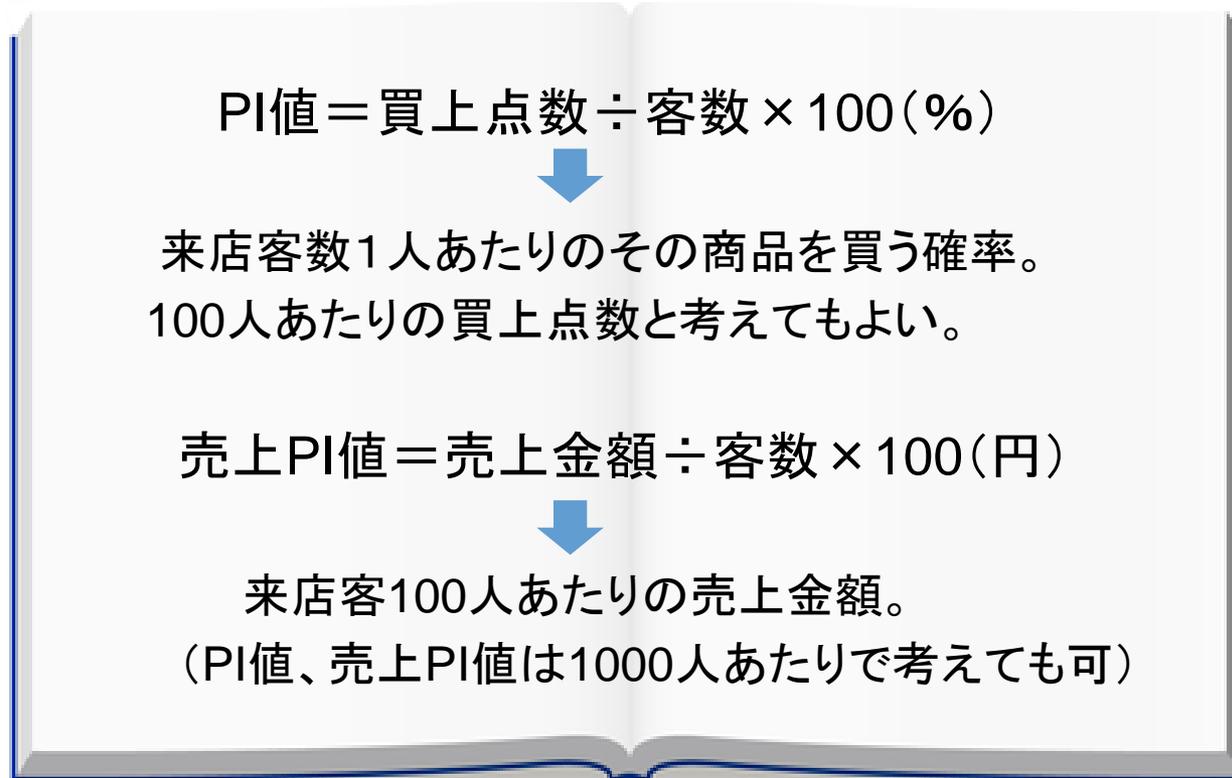


図5-3

「そうか、フェアで来店客数が増えれば売上が増えるのはあたり前か。PI値は来店客数の増加による売上アップ分をとるのか。確かにトマトソルトを置かなくてもトマトが1.5倍売れたかもしれないよな」

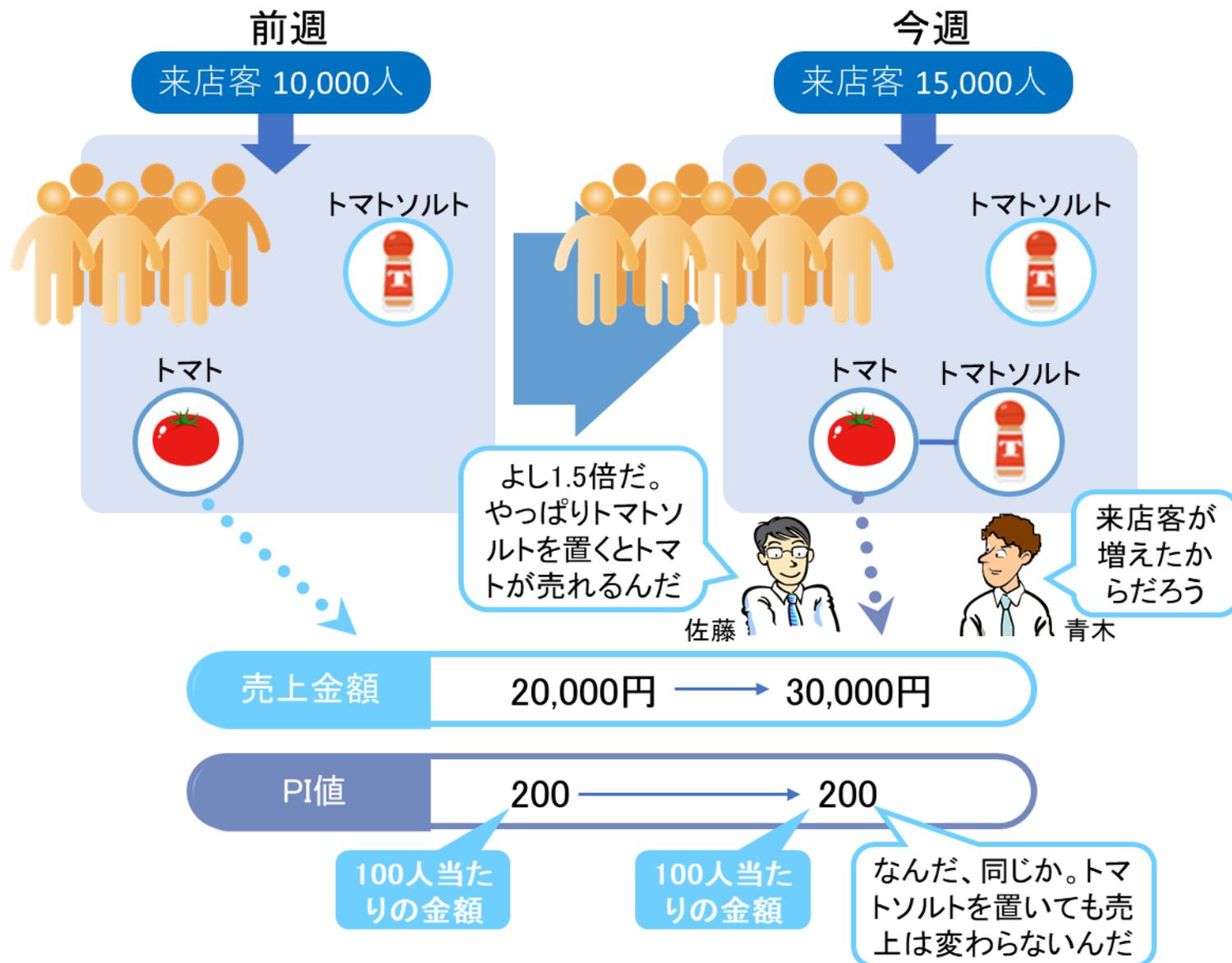


図5-4

佐藤は東バイヤーにレジ通過客数*を調べてもらい、売上PI値を計算した。

「よかった売上PI値でもあきらかにトマトの売上は伸びているぞ」(図5-5, [5-6](#))

売上PI値

	週	月	火	水	木	金	土	日	合計	前週比
トマト	10/6~12	210	178	186	169	226	210	194	1,373	1.29
	前週	198	122	187	142	98	162	152	1,061	—
トマト	10/6~12	50	66	78	45	89	106	72	506	5.56
ソルト	前週	19	25	0	9	8	4	26	91	—

図5-5

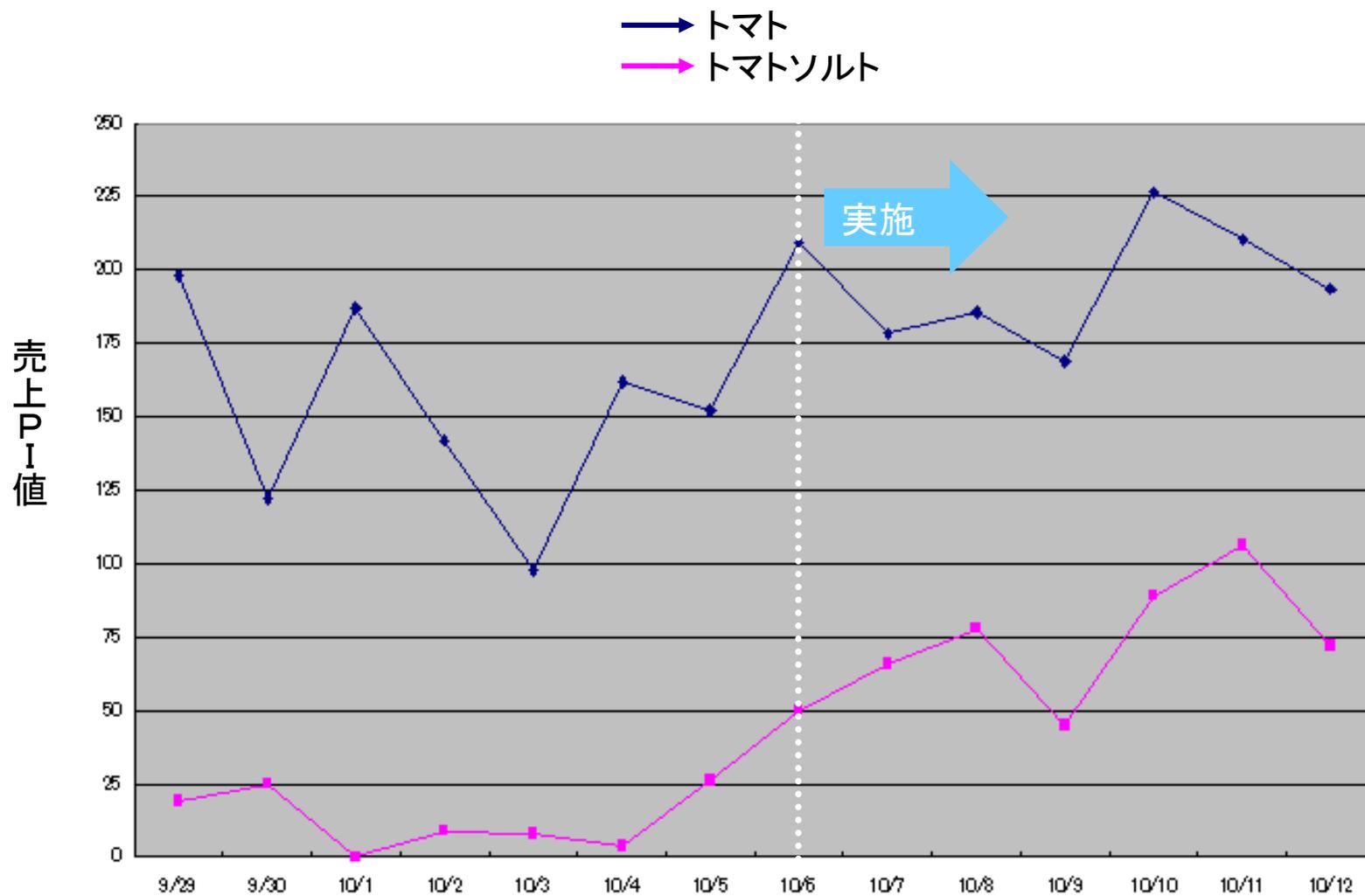


図5-6

佐藤 「売上PI値に直しました。それでもあきらかにトマトの売上は伸びています」

青木 「何か直感的にはこの数字、少し出すぎていると思うんだけど。わきにトマトソルトを置いただけでトマトがそんなに伸びるなんてちょっと考えづらいなあ。トマトが1.3倍か…。トマトソルトは2ヶ所陳列*だろうから伸びるのは当然としても…」

佐藤 「そうですか？」

青木 「これって実施前後で何か価格とか販促*とか変えていない？」

佐藤 「少し価格を下げました。それとトマトソルトの試食をマネキンがやりました」

青木 「それじゃトマトソルトを置いた効果がわからないじゃない。トマトソルトを置かなくても価格を下げればトマトが同じだけ売れたかもしれないでしょう。大体メーカーのセールスは実験、実験といいながら、できる限りの良い結果を作っているんだよな。だから他の店でやった結果は信用しないんだ」

佐藤 「もう少し頭を整理してみます」

「基本に戻ってよく考えなきゃ。まず母集団はなんだ。うーん、MCスーパー中央店のトマトの売上PI値か？ちがう。MCスーパー全体か。少し狭いなあ。だったら本社の成功事例発表会に出す必要はなくなる。そうか、MCスーパー中央店とよく似た青果売場を持った店舗のトマトの売上PI値か」

佐藤は絵を書いた。 (図5-7)

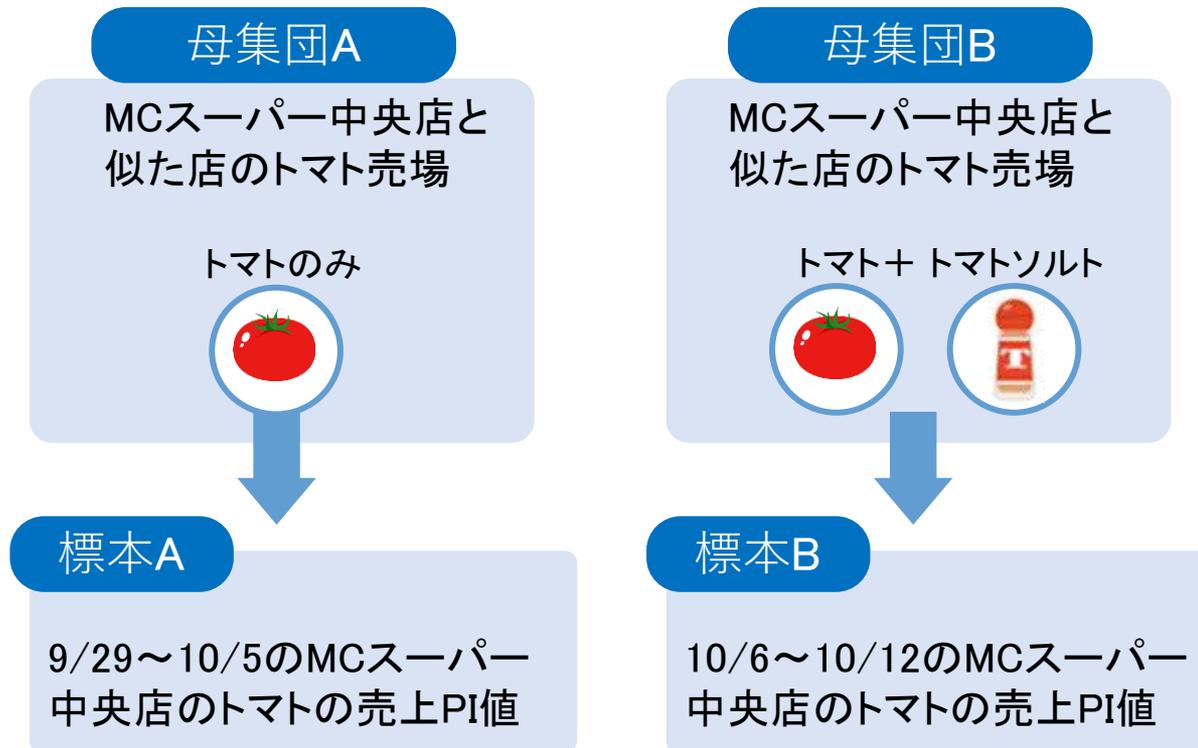


図5-7

「トマトソルトを置かないときの母集団と、置いたときの母集団は別の母集団のはずだ。これをAとBと考えよう。標本はAが9/29~10/5、Bが10/6~10/12のMCスーパー中央店のトマトの売上PI値か。トマトソルトが『効果なし』となるときはAとBは同じ母集団か。逆に考えれば、別の母集団といえれば『効果あり』となるのか」

「この標本を見て、母集団Aと母集団Bがちがうものだといえればいいんだな。でもちがうに決まってる。Bでは価格も下げて、試食もやっているから、標本Aと標本Bはトマトソルトを置いても置かなくてもまるでちがう母集団だ。このちがいがトマトソルトのせいかな、価格のせいかな、試食のせいかわからなくなったということか。ということは他は何もいじらないで、トマトのわきにトマトソルトを置けばよいのか。青木さんってやっぱり論理的だなあ」

翌日佐藤は青木を訪ねた。

佐藤「私はトマトソルトをわきに置けばトマトの売上は伸びると考えていますが、中央店の実績ではこれが実証できませんでした。是非、青木さんの所でやらせてください。」

青木「やっぱりこの結果を見せられるとやってみたい気もするよな。やってみるか。やるんなら1週間じゃデータが少ないので、来月の頭の週から1ヶ月やろう。それで今月と比較しよう。まあ、今月も来月もそんなに消費者の購買行動は変わらないだろう。価格も同じで、一切の販促なしで、トマトソルトとのクロスマーチャンドライジングだけをやって、結果を見よう。これがうまくいけば、他の野菜とのクロスマーチャンドライジングもやってみよう」

実験は実施され、1ヵ月後佐藤は青木よりPOSデータと来店客数をもらい、週および月単位の売上PI値を計算した。(図5-8、[5-9](#))

売上PI値(週単位)

	10月第1週	10月第2週	10月第3週	10月第4週	10月
トマト	135	130	125	119	127
トマト ソルト	12	15	11	14	13
	11月第1週	11月第2週	11月第3週	11月第4週	11月
トマト	129	131	139	129	132
トマト ソルト	26	36	28	33	31

- 週の売上PI値はその週の来店客数（レジ通過客数）と売上額で計算
- 月の売上PI値はその月の来店客数（レジ通過客数）と売上額で計算

図5-8

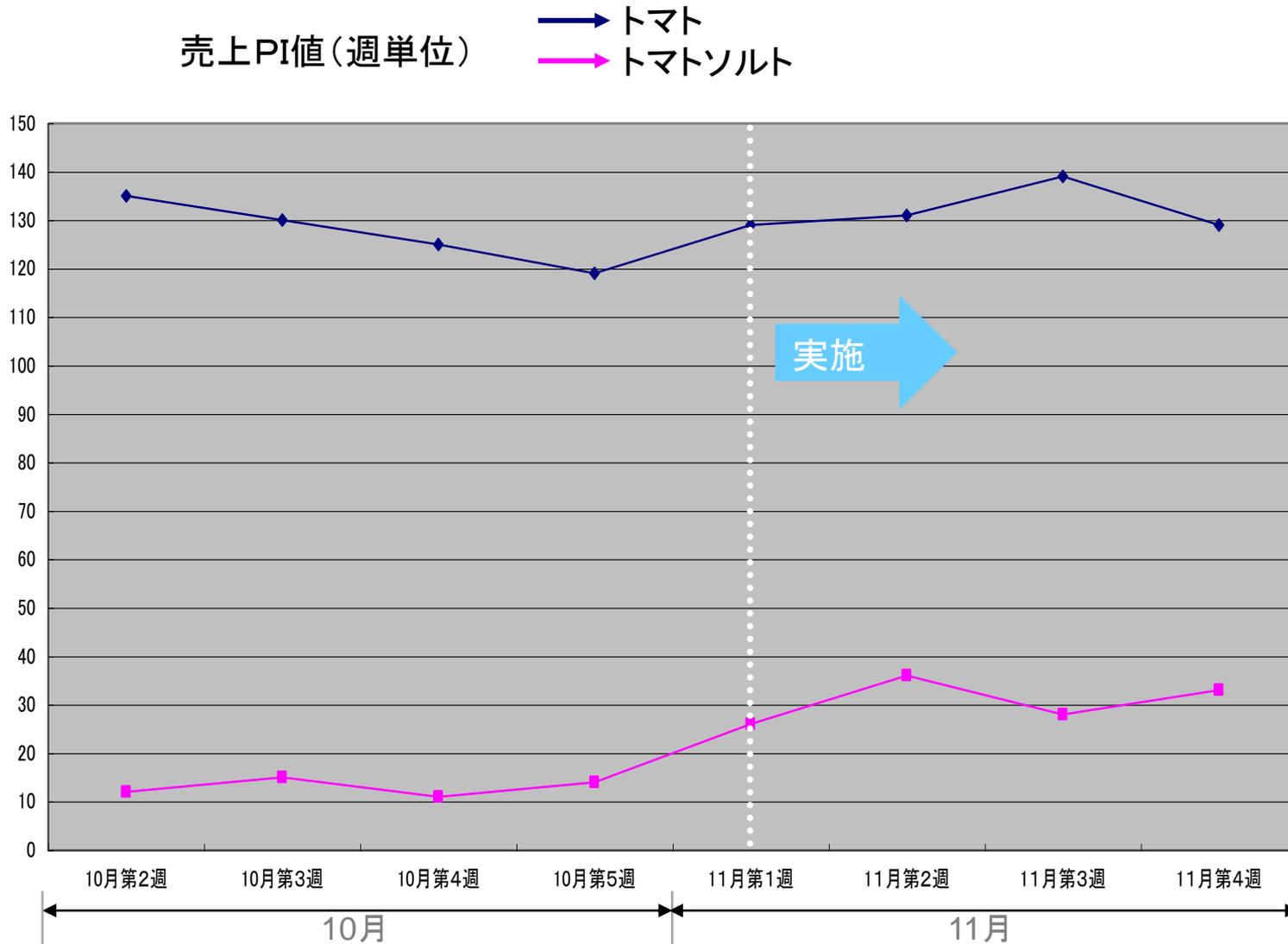


図5-9

青木「まあトマトソルトは当然として、トマトも明らかに上がっているよな。トマトの月の売上PI値は127が132か。わりとうまくいったんじゃないか。意外に効果があるんだね」

佐藤は中村支店長の言葉を思い出し、実験をもう1店やってから本社の「セールス成功事例会」へ出すことにした。

佐藤「青木さん、もう1店で同じことをやってみたいんですが、別の店のバイヤー紹介してもらえませんか？」

青木「それならうちの本部へ話を持ち込めよ。本部にいる高橋は私の同期だから彼を紹介するよ」

佐藤は実績データを持ってMCスーパー本部の高橋を訪問した。しかし高橋の反応は意外と冷たいものであった。

高橋「メーカーからよくこういう話は持ち込まれるんだよね。本当に効果はあるのかなあ。生鮮に加工をクロスすると売場が混乱して収拾つかなくなっていくんだよね。ところでトマトの売上PI値はどの位ちがうの？」

佐藤「クロス実施前の10月の売上PI値が127、で実施後は132です」

高橋「結果は増えているけど、たまたまじゃないの。もう少し様子みたいね」

帰り道、佐藤は思った。

「『効果が出た』と言い切れないのはくやしいよなあ。こっちが言い張っても水かけ論だし。でもこんなことやっているのと、仮に効果が出ても、いつも『たまたま』で終わってしまうよな。長くやってるうちにPI値が少しでもブレたりしたら、もう少しやってみようとなっていて、いつまでたっても効果があることを証明できず、結局1店でしか実施できないことになってしまう。そのうえ他のクロスマーチャンダイシングの提案もできないし、せっかくのビジネスチャンスを逃しているよなあ。でもこんなこと皆が悩んだはずだ」

佐藤は土屋の言葉を思い出していた。

佐藤は帰社後、土屋の所へ駆けつけた。

佐藤「土屋さん、悔しいです。どうしてもお客様を説得できません」

土屋「どうした」

佐藤は経緯を説明した。

土屋「実験前後での環境をそろえたのはよかったな。ただ青木さんも佐藤も実験当事者で、心の中で『うまく行け』と思っていたよな。これが2人の共通の思いだ。でもその結果を見た高橋さんは冷静だよな。確かにたまたまかもしれないし、効果あったのかもしれないよな」

佐藤「でも『たまたま』と『効果ある』はどこに境があるのですか？過去の賢い人はどう考えたんですか？」

土屋「いい所に気づいたね。過去の人々の知恵を使って何とか冷静な高橋さんを論理的に説得しようと思うことが大切だ。これが検定だ。検定は推定の裏返しだ。だから検定を知るには推定の考え方を整理しておかないと理解できない。推定には大きく分けると点推定と区間推定があるんだ。まず点推定と区間推定について勉強してみて」

佐藤は「誰でもわかる統計の本」で推定について調べた。(図5-10)

点推定

母集団で推定したい値を、標本を用いた1つの値で推定するものをいう。標本の平均値をもって、母集団の平均を推定するのが、その典型的な例である。

区間推定

母集団での値が、ある確率の範囲内で、一定区間内に入っていることを示すものを区間推定という。例えば正規分布に従う母集団の平均 μ を推定するとき、 n 個の標本から作られた平均値を、 m 標準偏差を σ とすると、 μ は95%の確率で

$$m - 1.96 \times \frac{\sigma}{\sqrt{n}} \leq \mu \leq m + 1.96 \times \frac{\sigma}{\sqrt{n}}$$

にある。この区間を信頼区間、95%を信頼度という。信頼度99%では1.96が2.58となる。

図5-10

「点推定の意味はわかった。区間推定は要するにズバリの平均値ではなく、いくつかからいくつかに入っている確率が95%と考えるわけだ。なんとなく数学らしいなあ。でもこの式の意味がわからない」

翌日佐藤は土屋に質問した。

土屋「この式？これは前に似たようなものをやったろう。覚えてないか」

佐藤「わかった、安全在庫ですね」

土屋「そうだ。安全在庫でやった安全係数にあたるのがここに書いてある1.96であり、2.58だ。かけている相手は標準偏差と思えばいい」

佐藤「そうか。わかった。例えば標本の平均値100、標本数が100、標準偏差が10の時、この計算をして母集団の平均は95%の確率で98から102にあるとやるんですね」

土屋「そうだ、ここまでわかったらいいよいよ検定にいこう」

土屋 「まず対象となる母集団から考えよう。何だと思う」

佐藤 「それは前に1人で考えたんです。どうも2つの母集団があると思います。1つはMCスーパーと似たような店舗でトマトのみを置いたトマト売り場の売上PI値です。もう1つはやはりMCスーパーと似たような店舗でトマトのわきにトマトソルトを置いたトマト売り場の売上PI値です」

土屋 「だんだん数学を理解して表現がしっかりしてきたな。まあわからない人が聞くと、ずい分理屈っぽいと思うけどな」

佐藤 「今はやりのロジカルシンキングと言ってください」

土屋 「ここから先は本を読んでも、佐藤にはきっとわからないから、一緒にやっ
ていこう」

佐藤 「助かります」

土屋 「まず、検定の方針だ。『トマト売り場にトマトソルトを置いても効果がない』と考えよう」

佐藤 「えっ、逆を仮定するんですか？」

土屋「数学ではよくやる手だ。直接は証明できないことがあったとき、その仮説の反対を仮定して、矛盾がないかをチェックするんだ。そこで明らかに矛盾があれば『反対の仮説』が間違っていたとなり、もとの仮説が正しいと考えるんだ。『Aではない』が明らかに間違っている時は『Aである』が正しいと考えるんだ。数学でいう背理法だ。聞いたことがあるだろう。」

佐藤「ありません」

土屋「しょうがないな。『トマトソルトを置いても効果がない』という仮説を立て、標本の統計量を見るんだ。そして標本の統計量はその値をとる確率を出すんだ。さっきの区間推定では95%の確率で入る範囲を出しただろう。今度は逆にその値になる確率を出すんだ。そしてその確率が一定値より小さければ、仮説が誤っていたと考えるんだ。この時仮説は誤りとして捨てることが多いんで、帰無仮説というんだ。そして捨てることを数学ではかっこよく棄却っていうんだ」

佐藤「どうして確率が小さいと誤ってるんですか？」

土屋「本当に効果があったとしても『たまたま効果がない』ように見えることもあるよな。だからどんなに効果があってもなくても、その値にならないという結論は出せない。つまりその値をとる確率が小さいか大きいからだ。そして、一定の確率よりも小さいということは、そんなことはめったに起きないということだろう。めったに起きないことがたまたま起きるなんておかしいから仮説は誤りと考えたほうが良いと考えるんだ」

佐藤「今回のトマトの場合でもう1回説明して下さい」

土屋「『トマトソルトを置いても効果がない』ということが本当だとして、今回のような結果、つまりPI値が127から132に増える確率を出すんだ。もしそれが1%なら100回に1回しかないということだ。100回に1回しかないことが、MCスーパーの中央店の売場で『たまたま』起きたと考えるよりも、そもそも『効果がない』と仮定した事がまちがいと考えるべきだろう」

佐藤「でも効果がなくても売上PI値は127から132になることはあるんでしょ」

土屋「もちろんある。さっきの例だと1%だ。100回に1回はあるんだ。これを難しくいうと第1種の誤りというんだ。つまりそれでもトマトソルトに効果がないことは1%の確率でありうるということなのさ」(図5-11)

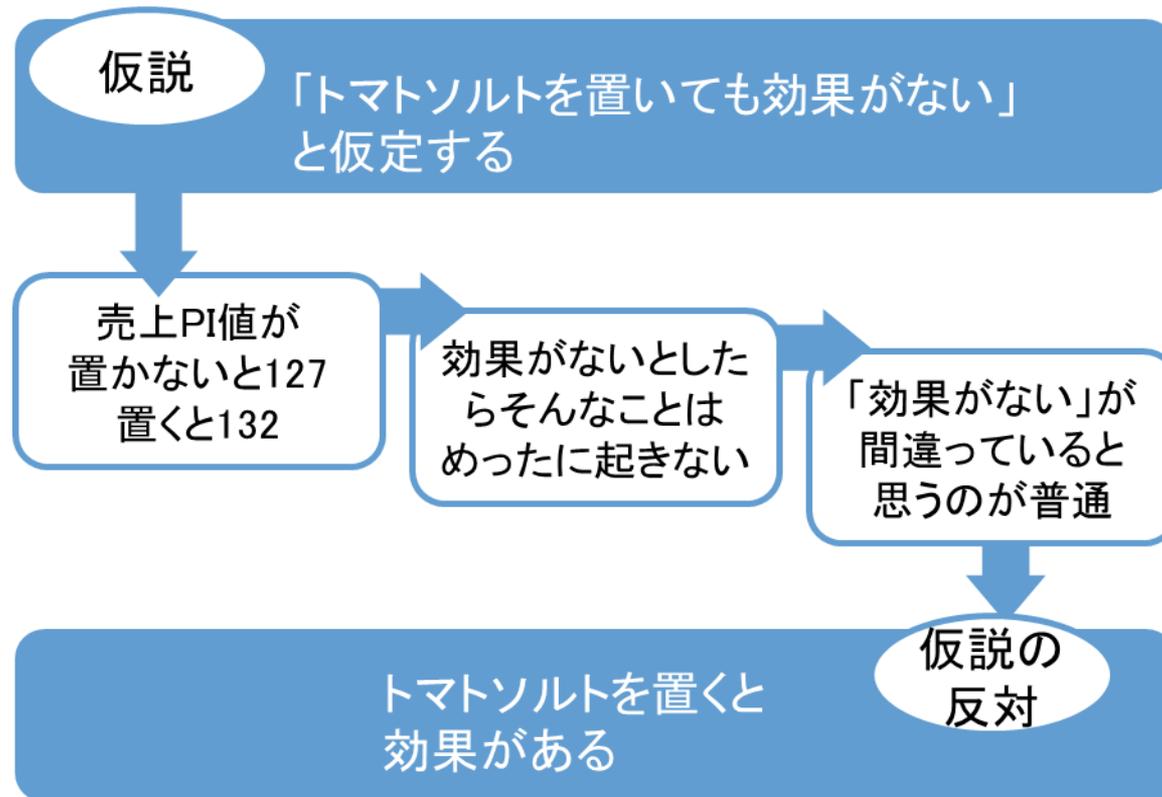


図5-11

土屋 「じゃあ、データは2つに分けよう。トマトソルトを置かなかった方がA、置いた方がBだ」

佐藤 「Aは10月、Bは11月です。」

土屋 「データは多い方がいいから1日ごとの売上PI値にしよう。AとBの標本数はいくつだ」

佐藤 「A、Bとも4週28日分でそれぞれ28個です」

土屋 「ここでまずトマトソルトを置いても効果がないとしよう。トマトソルトを置いても置かなくても効果がないんだからAの28個、Bの28個とも同じ母集団からの標本と考えられる。まさか今回はトマトソルト以外いじってないよな」

佐藤 「今回はいじってません」

土屋 「まずA、Bそれぞれ28個の標本値の平均値を出してみろ。小数第1位でいい」

佐藤 「もう出てます。Aが127.4で、Bが132.3です」

土屋 「似たようなもんだな」

佐藤 「やめてください。そんなこと言うの」

土屋 「今回の仮説である『トマトソルトを置いても効果がない』というのは、『AとBは同じ母集団であり、平均値には差がない』と言う仮説と同じことだ。つまり、この127.4と132.3は似たようなもので、たまたまちよっとちがっただけと言う仮説だ。これをひっくり返すんだ。本などでは平均値の差の検定と書いてある。」

佐藤 「確かに同じ母集団から標本を取り出しても平均値は少しちがっていてあたりまえですよ。それからどうするんですか」

土屋 「同じ母集団から28個ずつ標本を2回とって、佐藤の手元にあるような56個の数字になる確率を出すんだ。それが一定以下ならこの仮説は誤りとしよう。この確率を有意水準というんだ。」

佐藤 「有意水準はいくつですか。」

土屋 「区切りの良い所で5%か1%あたりが多いな。いくつにする？」

佐藤 「じゃ5%で…」

佐藤 「ところでその56個になる確率はどうやって出すんですか？」

土屋 「さっきやった区間推定の逆だろう」

佐藤 「区間推定は95%の確率で、平均が一定の範囲内に入っていることを考えた。今度はこういう平均になる確率が何%かと考えれば良いのか。やり方は浮かばないけどできそうですね」

土屋 「それが大切だ。やり方よりも発想だ。こういう2つの平均値の問題では t 値というのを使う。 t 値というのはさっきのAとBのそれぞれの標本値から計算するんだ。ただここではAとBのバラツキには差がないという前提が必要なんだが…。まあこの辺はあまり深く考えるな。いずれにしてもAとBの56個のデータから機械的に t 値というのを計算すればいいんだ。この t 値の大小で確率がわかるんだ。こういったものを検定統計量というんだ。まず5%の場合の t 値がいくつになるか、表で見よう」

佐藤 「表ってどれですか」

t分布表

有意水準 自由度	0.01	0.02	0.05	0.10
1	63.657	31.821	12.706	6.314
2	9.925	6.965	4.303	2.920
3	5.841	4.541	3.183	2.353
4	4.604	3.747	2.776	2.132
5	4.032	3.365	2.571	2.015
6	3.707	3.143	2.447	1.943
7	3.500	2.998	2.365	1.895
8	3.355	2.896	2.306	1.860
9	3.250	2.821	2.262	1.833
10	3.169	2.764	2.228	1.812
...
26	2.779	2.479	2.052	1.703
27	2.771	2.473	2.052	1.703
28	2.763	2.467	2.048	1.701
29	2.756	2.462	2.045	1.699
30	2.750	2.457	2.042	1.697
...
40	2.704	2.423	2.021	1.684
60	2.660	2.39	2.000	1.671

図5-12

土屋「本に t 分布表というのがついているだろう。これを使えばさっきやった有意水準 5 % と自由度から求められる。自由度というのは各々の標本から 1 を引いたものだ。28 個ずつの標本だから 27 と 27 で 54 だ。54 で 0.05 の t 値はと…。この表には自由度 40 と 60 しかないが、まあいいや。どっちも似たようなもので、2.0 だ。この 2.0 が限界値だ。 t 値がこれより大きくなる確率は 5 % もないんだ。2.0 より大きいかわ小さいかを見よう。56 個のデータで値を計算してみろ」

佐藤「式を教えてください。」

土屋「56 個のデータは表計算ソフトに入っているだろう。そのソフトが式を覚えているから大丈夫だよ」

佐藤「えーと、データ分析ツールを選んでと。あれ 検定ってたくさんありますよ」

土屋「さっき A と B のバラツキが等しいと仮定したとあったな。『 t 検定の等分散を仮定』というのを選べ」

佐藤「入力範囲を入れてと。これで OK ですか」

土屋「OK だ」

佐藤 「もう出ました。 t 値を見ればいいんですね。あれ、 -3.0 と出ました」

土屋 「今の仮説は『AとBの平均値に差がない』というものだったな。これは『A-B』の平均値が0と考えても良いかというのと同じだ。これを計算しているのが t 値だ。Aの方が小さければ当然マイナスになる。この時はマイナスを取れば良いんだ」

佐藤 「ということは3.0か。やったあ！2.0より大きいから、こんなに大きくなることは確率5%もない。だから仮説の誤り」

土屋 「これでAとBに差がありといえる。これをよく有意差というんだ」

佐藤 「ところで、この2.0は表計算ソフトではどこに出ているのですか」

土屋 「それは境界値というところだ。」

佐藤 「あれ、境界値に片側と両側がありますよ」

土屋 「なかなか鋭くなってきたな。 t 値というのはさっきのA-Bのことだ。このA-Bが t 分布というグラフになるんだ」

と言って土屋はグラフを紙に書いた。

佐藤「正規分布とそっくりですね」

土屋「自由度が大きくなると、つまり標本数が大きくなるとほとんど正規分布になるんだ。だから佐藤は正規分布の一種で考えてもいいぞ。この t というのは『 $A-B$ 』の確率分布と考えるんだ。 $A-B$ の平均が a より大きな所がここ、 $-a$ より小さい所がここだ」

とって土屋は斜線を引いた。(図5-13)

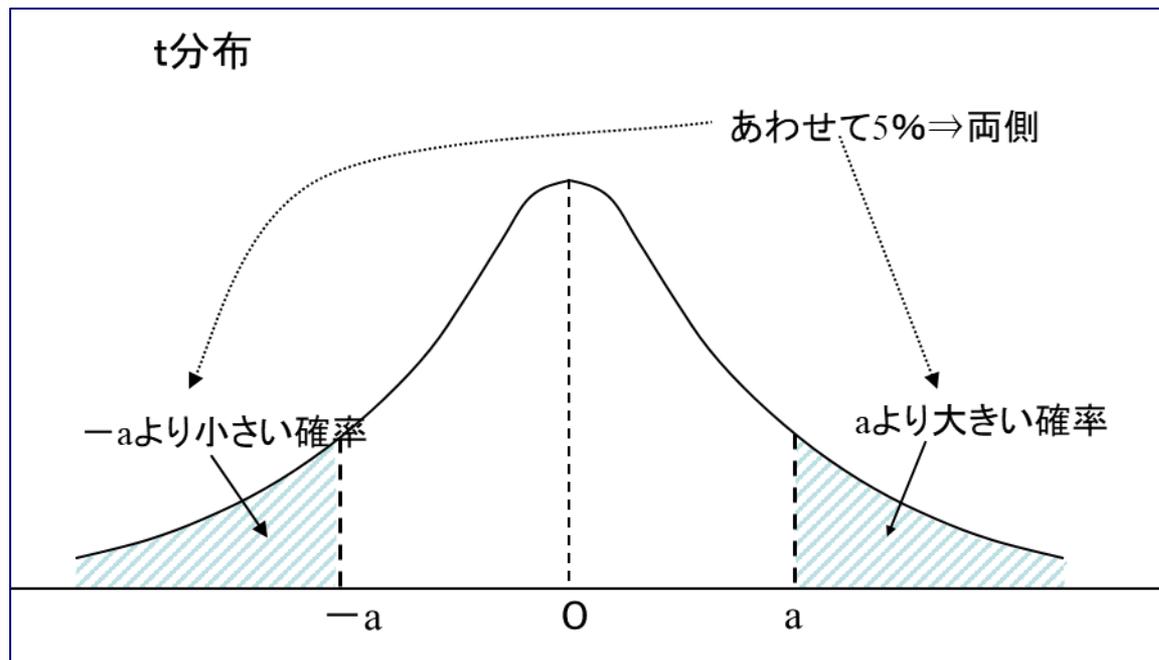


図5-13

土屋「両側というのはこの2つの面積を足したものが5%になる所が。aにあたるのがパソコンに出ている両側境界値2.0だ。つまりAよりBが大きい方がaで、AよりBが小さい方が-aだ。片側というのは、Aの方がBより小さくなるなんてありえない、つまりA-Bは絶対にプラスになると考えられる場合だ。そう考えるとこうなる」

と言って図5-14のように斜線を引いた。

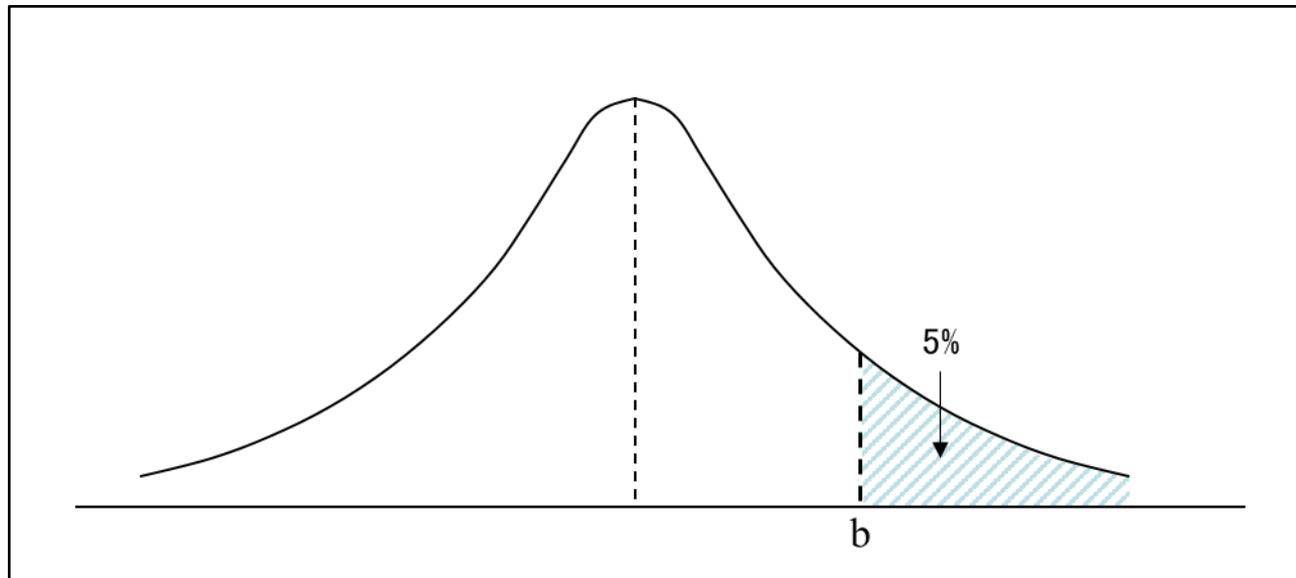


図5-14

土屋 「この b にあたるのが片側境界値1.6だ。Aをトマトソルトを置いた場合、Bを置かない場合で考えてみる」

佐藤 「両側というのはBがAより大きくなることもあると考えるわけだから、トマトソルトを置くとトマトの売上が落ちないとは言い切れないから今回は両側でやった方が説明しやすいな。両側で考えても2.0より t 値は大きいので『効果なし』は誤りだ」

土屋 「そうやるんだ。両側がいいか片側がいいかなんてトマトソルトやその売場を知っている人以外に決められない。そうやって人間が知識を使って数字を選ぶんだ。コンピュータはただ数学者が考えた式どおりに計算しているだけだ。自分で知識を使って数字を選べばお客様にだって説明できる。そしてお客様の意見を聞くこともできる」

佐藤 「よし『トマトソルトの効果はない』は誤り。よって『トマトソルトの効果あり』。つまり『トマトの売上が伸びる』だ。やったなあ。なんだか3段論法みたいでまわりくどいけど…。でも土屋さん、こんなややこしい方法しかないんですか？」

土屋 「あるかもしれない。でもまだこれ以上良いやり方を思いついた人は人類史上いない。君が考えればノーベル賞をもらえるよ」

佐藤 「これを使うしかないのか」

土屋 「この仮説検定の第一種の誤りは何%だ」

佐藤 「5%でしょ。つまりもとの仮説である『効果なし』が正しい確率でしょう。そういえば不思議なんですけど、第一種ということは第二種もあるんですか？」

土屋 「もちろんある。それは今の逆だ。つまり仮説が棄却されなかった時で、つまり t 値が2.0より小さくて『効果なし』と判断したが、本当は『効果ある』という誤りだ。」

佐藤 「さてお客様にプレゼンするんだから少し整理しておこう」 (図5-15)

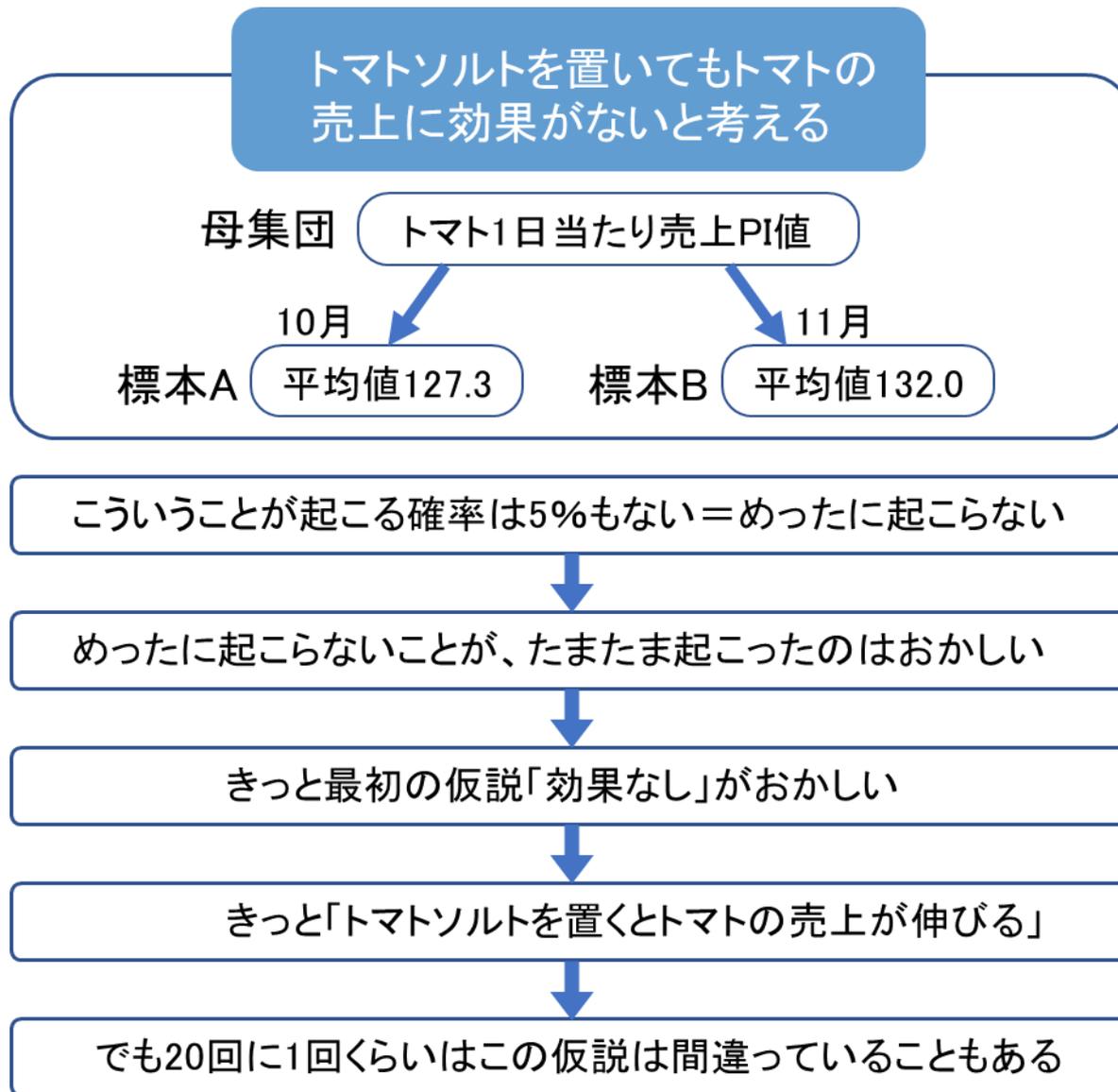


図5-15

佐藤は高橋への説明を無事終えた。

高橋「そうか、勉強になったよ。私も文系で今1つ検定の意味がわからなかったんだ。昔統計のプロの話セミナーで聞いたんだが、やり方はわかるんだが意味がわからなかったんだ。失礼だけど君のようなアマチュアから聞いたほうがよくわかったよ。

ところで話は全然ちがうんだけど、前から引っかかっていたことがあったんだ。うちの20店で、ある実験をやったんだ」

と言って、高橋は実験結果のレポートを佐藤に見せた。
(図5-16)

××カップ麺PI値

××カップ麺PI値			
	()内は前週比		
	第1週 (1フェース)	第2週 (2フェース)	第3週 (3フェース)
A店	4.3	8.2 (1.91)	5.6 (0.68)
B店	2.6	6.2 (2.38)	3.8 (0.61)
⋮	⋮	⋮	⋮
合計	90.2	178.6 (1.98)	124.0 (0.69)

高橋「実は今、私は店舗でのフェース効率の研究をしてるんだ。現場の店長、バイヤーたちはボリューム感*よりも品揃え感*を大切にしたいというんだ。つまり色々なものを少しずつ置きたいということだ。私の個人的意見では死に筋をどんどんカットして売れ筋に絞り込んで、その売れ筋のフェースを増やした方が効果的だと思ってるんだ。売れ筋のフェースを増やせば売上がもっといくんじゃないかということだ。そこで現場と一番もめているカップ麺について、売れ筋上位5アイテムを使って20店で実験させたんだ。この5アイテムを1週目は1フェースとり、2週目はすぐとなりに同じ商品でもう1フェースとる。3週目はこれをもとの1フェースに戻すという形だ。これがその結果なんだ」

佐藤「あれ前週比が皆よく似てますね」

高橋「もちろんイレギュラーな値もあるけど、5アイテムの合計を見てみると驚くべきことに1フェースを2フェースにするとPI値は約2倍に、2フェースを1フェースに落としたとしても、翌週だけ見ると70%くらいにしか落ちないんだ」

佐藤 「これって不思議ですね。1フェースを2フェースにしたくらいで売上が2倍になるのもちょっと予想外ですが、2フェースを1フェースにして70%も出るというのは何か顧客に認知効果が残るんですかね」

高橋 「ただ各店の店長が、たまたまこうなっただけじゃないかというんで困ってたんだ。そこで君から今話を聞いて、検定を思いついたんだ。どうしたら良いだろう」

佐藤 「これって今回と同じだと思います。まず第1週から第2週ですが、これは1フェースあたりのPI値を見れば良いと思います。つまり第2週を全部2で割ればよいわけです。それで第1週と第2週について今回のように平均値の検定をやれば良いと思います。『フェースあたりのPI値は等しい』という仮説を立て、棄却されれば『2倍になるとはいえない』、棄却されなければ『2倍になる』といえるところを考えます」

高橋 「棄却されなければ『2倍になる』と考えて良いということか」

佐藤 「もう1つの第3週で元に戻す方ですが、第2週に0.7をかけてという手もあるんですが、むしろこっちも1フェースあたりにして考えて『第2週と第3週は等しい』と仮定するほうが良いと思います。こちらはきっとこれが棄却されますから『等しくない』となりますよね。そうすると『2フェースを1フェースにしてもすぐにはPI値が半分になるわけではない』といえるようになります。ここまでではないですかね」

高橋 「よしやってみよう。表計算ソフトで出来るんだったよね」



ボリューム感

店舗で1つの商品をたくさん置くことで、顧客に安さをイメージさせる陳列。

品揃え感

店舗に色々なものを置くことで、顧客に品種の豊富さをイメージさせる陳列。

高橋 「もう1つだけ相談にのって。今我々は店舗のフェーシング*とともにレイアウト*全体を見直すことも考えているんだ。そこで客動線*を一生懸命調べてるんだが、まずお客様は店の入口から入ってどっちへ行くかを知りたいんだ。これによってレイアウトの作り方がちがうんだ。そこで、ある店でお客様が、店に入ってまず右に行くか、左に行くかを実際にカウントしてみたんだ。その結果がこれだ」

右へ行った人	73人
左へ行った人	51人
合計	124人

高橋 「私はどう見ても右へ行く人が多いんで、多くの人間はまず右へ行くと言いたいんだ。でも社長はたまたまじゃないかと言うんだ。これじゃあ調査人数を増やしていっても、ずっとたまたまで終わるんじゃないかって心配になるんだ。とりあえず答えを出したいんだけど、どうしたら良い？」

佐藤 「ちょっと考えてみます」



フェーシング、レイアウト

大分類商品（生鮮食品で言えば鮮魚、青果など）を店舗のどこに配置していくかを定めることをレイアウトと言う。中分類ごとやカテゴリー単位でどこに商品を置くかをゾーニングと言い、各商品のフェースを決めることをフェーシングと言う。

客動線

顧客が店舗に入ってからどのように動くかを言う。単に動線とも言う。

佐藤は帰社して考えた。

「うーん、これって平均値ってないよな。標本が1人増えても73人か51人に1人増えるだけだし…。バラツキもないし…。うーん、わからない。土屋さんに聞くばかりでもさみしいし、よし家に帰ってもう1回本を読もう」

さまざまな検定

- ・母平均の検定→正規母集団の母平均を検定する。
- ・バラツキの検定→2組の母集団の分散の有意差検定。
 F 値を使うので F 検定ともいう。
- ・平均値の検定→2組の母集団の平均の有意差検定。
 t 値を使うので t 検定ともいう。
- ・相関の検定→ 2組の母集団の相関関係の検定。
- ・適合の検定→ 理論値に適合しているかを検定。 χ^2 値を使うので χ^2 検定ともいう。

図5-18

「へーこんなにあるんだ。母平均の検定か。帰無仮説は『母平均と予測平均は等しい』か。これを棄却すれば『等しくない』となるのか。あれっ、この人は『等しくない』といたいのか。自分が考えたものと反対のものを想定するんだっただよな。『等しくない』じゃその平均が使えないということだろう。変だなあ。しょうがない土屋さんに電話で聞いてみよう」

土屋 「よく勉強してるなあ。帰無仮説というのは『なくす』という意味だから、本当は自分の仮説とは反対のほうが良いんだ。でも母平均の検定の場合、『等しい』として仮定すると数字をいろいろ加工できるけど、『等しくない』では加工できないだろう。だから『等しい』と仮説を置くしかない」

佐藤 「そうすると『等しい』という仮説が棄却されないと、『等しい』となるんですか？」

土屋 「帰無仮説が棄却されない時は、正確にいうと証拠不十分で『否定できない』という感じだなあ。『等しいという仮説が誤っているということが今回のデータではいえない』というのが数学的ニュアンスだ。ふつうの表現では要するに『等しい』という結論だ」

佐藤「そういえばさっきの『2フェースで2倍』もそうだった。棄却されないということはまあ『2倍になる』と考えていいということか」

土屋「ここで本当は2倍でなければ第2種の誤りだ。さっきやったな」

佐藤は納得して電話を切った。

「次を見よう。バラツキの検定か。これは標準偏差のちがいを見るんだな。あれ分散て何だ。そうだ標準偏差の2乗だったっけ。これは前にやった平均値の検定とそっくりだ。F値というのを使うのか。そうだ前のスーパーグリーンとナショナルヘルスハイパーでやってみよう。まず仮説は『ナショナルヘルスとスーパーグリーンのバラツキは差がない』だ。たしかまだ表計算ソフトにデータが入っていたな」

佐藤は前にやったスーパーグリーンとナショナルヘルスハイパーのデータをコンピュータのファイルで探した。「F検定を選んで、範囲を指定して、棄却率を0.05とセットしてと、あっ、もう出たぞ」

コンピュータの画面に検定の結果が出力された。

「検定統計量はこの『観測された分散比』か。ナショナルヘルスとスーパーグリーンの分散の比を見てるのか。そうか、これが1に近ければバラツキは等しく、大きければバラツキは等しくないんだ。これがt値にあたるF値のことだな。えーっと大体4.8だ。境界値は1.4か。ということはこの仮説は棄却だ。つまり『バラツキは等しい』という仮説は棄却されて、『バラツキに差があり』だ。当然だ。スーパーグリーンの方がコンスタントなんだから」

「そういえば前にトマトソルトの時、土屋さんが10月と11月の母集団のバラツキに差がないことが前提といかいっていたなあ。よし土屋さんに聞いてみよう」

佐藤は土屋に再度電話した。

佐藤「たびたびすいません。以前にトマトソルトの検定やったときに、土屋さんがちらっと、これはバラツキに差がないことが前提と行ってましたよね。どういう意味ですか？」

土屋「わかった、今度は分散の検定をやってるんだな。がんばってるな。これはもし10月と11月の標本を見て、平均値に差があったとしても、それはもともと母集団のバラツキが大きいのかもしれないからだ」

佐藤「バラツキが大きいと同じ母集団でも平均値に差が出ることもあるということか。トマトの売上PI値が毎日大きく変動していれば、10月と11月で平均が変わってもトマトソルトの効果とはいえないということか」

土屋「だから本当は平均値の差の検定の前にバラツキ検定をやらなければいけないんだ。バラツキの差がなく、かつ平均値に差があれば本当に『トマトソルトは効果あり』だ」

佐藤「じゃやり直しですか？」

土屋「心配するな。俺が代わりにやっておいた。10月と11月の標本値のバラツキに差はなした。正確に言えば『バラツキに差があるという仮説は、有意水準5%で棄却される』」

佐藤「よかった。勉強のため自分でも F 値を計算してやってみます」

佐藤は前にやったトマトソルトの売上PI値のデータで F 検定をやってみた。

「おっとこんなことをやっている場合じゃない。平均値の検定は前にやったやつだ。次の相関の検定は相関があるかどうかを見るんだろうな。こうしてみると今までやってきたことと同じ項目じゃないか」

「そうか今までやっていたのが推定で、その裏側にいつも検定があるんだ。推定して、本当にその推定でいいかを検定するのか。次の適合の検定は何かいけそうだぞ。これってカイ2乗と読むのか。エックスみたいだな。何々『この検定は計数値の検定に向いている。計数値とは度数を数えるようなものであり、計量値とはその数値を測るものをいう』か。今回は右へ行く人、左へ行く人を数えているんだから、ぴったりじゃないか。これで行こう。この検定統計量は、...」

(図5-19)

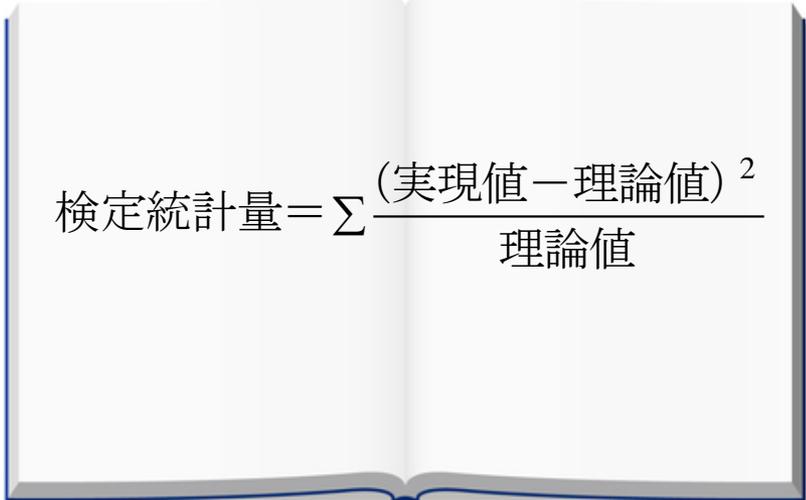

$$\text{検定統計量} = \sum \frac{(\text{実現値} - \text{理論値})^2}{\text{理論値}}$$

図5-19

「この検定統計量はカイ2乗分布に従うか。よしできそうだ。まず仮説を立てよう。土屋さんにいわれたように『等しい』方が良さそうだから、『人間は右回りでも左回りでもなく、気分次第で右へいたり左へいたりする』だ。そうすると理論値は全体の半分だ。棄却率は5%、と」

「これでまず右へ行く人の実現値から理論値を引いて、2乗して理論値でわって、左へ行く人も同様にして足すと、3.9か。これでカイ2乗分布表の5%の所を見てと。自由度は右と左で2つのため、2から1ひいて1か。3.84か。ということはギリギリで棄却か。『等しくない』つまり『右へ行く傾向にある』か」

翌日佐藤は高橋に説明した。

「今回のデータをカイ2乗検定で分析したところ、『右へ行く傾向にある』と考えられます。というよりも『右へも左へも同じように行くとはいいがたい』となります。ただし棄却率は5%ですので20回に1回は誤りがあります。またこれはこの124人のデータから考えるとそうなるということでデータ量が増えれば結論が変わる可能性もあります…」

	理論値	実現値
右へ行く	62人	73人
左へ行く	62人	51人

図5-20

佐藤は必要に迫られ、さまざまな統計的手法を学んだ。そして彼は今、二つのことを考えている。一つは、自らが学んだことをどうやってまわりの人たちに伝えていくか、ということである。「自分が学んだこの統計的ナレッジは今のミドリ食品の営業マンに求められているものじゃないかな。それを同僚や後輩に教えてあげるのも、自分の仕事だろう。まずは勉強会をやり、数字で悩んでいるケースを集めて、どうすればよいかを話し合っていこう。そしてこの結果を社内に公開していけばいい」

二つ目は、学んだ統計的手法をもっとさまざまな分野に活用することである。「この統計の考え方はじぶんがセールスとしてやってきた顧客との折衝だけでなく、予算、経営計画、管理会計などリーダーとしての仕事に使えるはずだ。真のリーダーは、自分の知識、経験、カンを活かして仮説を立て、数字をうまく使ってまわりの人に論理的に自分の仮説を説明できる人だと思う。よしがんばるぞ」

営業マンとしての現場の知識、ノウハウに加えて「数字を使う力」を身につけた佐藤は、ミドリ食品のリーダーへの一步を踏み出したと言えよう。

確率分布

最後に確率と統計の分野がたどってきた道を追いかけて、まとめてみましょう。もう1度第1章のコラムにある1等から4等のボールを思い出してください。ボールを引いて得る金額を x とし、この x をとる確率を $P(x)$ と表します。

$$P(x = 600) = P(x = 300) = P(x = 100) = P(x = 0) = \frac{1}{4}$$

この時、期待値を E とすると

$$E = 600 \times \frac{1}{4} + 300 \times \frac{1}{4} + 100 \times \frac{1}{4} + 0 \times \frac{1}{4} = 250$$

となることは第1章でもうやりました。

今度は在庫で考えましょう。 x が在庫量です。第4章で述べたように在庫は連続量として考えましょう。これもやったように x が a から b までの値を取る確率は

$$P(a \leq x \leq b) = \int_a^b f(x) dx$$

と表せます。この $f(x)$ が確率密度関数です。これもちらっとやりました。そしてこの $f(x)$ としてもっとも便利なものが正規分布となることもやりました。この時、上と同様に期待値 E を考えましょう。

x をとる確率は $f(x)$ ですので、上で $600 \times \frac{1}{4} + 300 \times \frac{1}{4} \dots$ と足していったのと同様に、期待値 E は $x \cdot f(x)$ をはてしなく足していくものといえます。

つまりとなります。

つまり $E = \int_{-\infty}^{\infty} xf(x)dx$ となります。

(∞ は無限に大きいもの、 $-\infty$ は無限に小さいもの。

つまりはじからはじまで積分するという感じです。)



このあたりからが、確率から統計の分野に入っていきます。統計の最初のテーマは標本から母集団の平均を推定する事です。最後なので少し数式を使ってみます。母集団の平均を μ (ミュー)、標準偏差を σ (シグマ) とし、標本の平均を m 、標準偏差を s とします。この時、95%の確率で

$$m - 1.96 \frac{\sigma}{\sqrt{n}} \leq \mu \leq m + 1.96 \frac{\sigma}{\sqrt{n}}$$

と推定することも本文でやりました。

\sqrt{n} で割るのは平均値 μ の標準偏差が $\frac{\sigma}{\sqrt{n}}$ になるからです。

しかしよく考えると σ は母集団の標準偏差であり、わかっていません（母集団の平均がわからないのに、標準偏差なんてわかりようがありません）。つまり μ を推定するためには先に σ を推定しなくてはならないのです。そこで仕方がないので、標本の標準偏差 s をそのまま σ の代わりに使おうと考えました。しかしそれではどうしても小さすぎてしまいます。つまりバラツキが小さく推定されすぎます（母集団からの標本が少ないとバラツキが小さくなってしまうのは何となくわかると思います）。

本来、分散 $s^2 = \frac{\sum(x_i - m)^2}{n}$ で計算するのですが、賢い人がこれを修正して

$s^2 = \frac{\sum(x_i - m)^2}{n-1}$ として少し大きくした方が、 σ をうまく予測できると証明

しました。本によっては分散、標準偏差の出し方が $n-1$ で割ってあるのはこのためです。

この $n-1$ で割ると思いついたのは、次のような理由からでした。個の標本なら分散は n で割れば良いけど、よく考えると σ を推定するのに m という計算を使っている。ということは、 m という平均値が決まって、さらに $n-1$ 個の標本値が決まれば、もう 1 つの標本値は自動的に決まってしまう。自由に値をとれるのは $n-1$ 個しかないと考えました。本文の「右へ行く」か「左へ行く」かの例で、全員で 124 人（右へ行く人と左へ行く人の平均 62 人）、右へ行く人が 73 人と決めれば右へ行く人は 51 人と決まってしまう。つまり自由に値をとれるのは「右へ行く」データだけです。この $n-1$ を自由度といいました。本文で自由度を標本から 1 引いたのはこのためです。ただ n がかなり大きくなってくると、 n も $n-1$ もほとんど変わらなくなってしまいますので、無視しても大丈夫です。

次にこの $m - 1.96 \frac{\sigma}{\sqrt{n}}$ の分布を調べ、これを t 分布（正確にいうと自由度 $n-1$ の分布）と名づけました。これが本文でやった t 分布の誕生です。標本が増えて自由度が上がってくると t 分布は正規分布に近づいてきます。自由度 30 をこえたあたりから、もう正規分布と考えても良いくらいになります。

したがって本文でやったように、標本数は十分大きければ正規分布表を使って平均を推定してもOKです。

カイ2乗分布というのは X が正規分布のとき、この x^2 の分布をいいます。適合の検定では（理論値－実測値）²を使うことを思い出してください。これがカイ2乗の2乗の意味です。さらに2つのカイ2乗分布の比を使うのが F 分布です。分散の検定で分散の比を使いましたが、分散の比は分子と分母に2乗があることを考えれば何となくわかると思います。

これで確率、統計の理論の90%は終わりです。ここまで読み終えたビジネスパーソンは確率、統計のプロといってもおかしくありません。後の10%は万が一数学者になる時のためにとっておきましょう。これであなたは卒業です。

最後に数学のプロである土屋さんが本文で述べていた格言を挙げておきます。もう一度通して読んでいただき、数学、予測の考え方を理解してください。

「統計は昨日の結果。それに今日の変化を加えて予測する。数字を加工するのはコンピュータ。今日の変化の反応は人間が考える」

「数字の見方で大切な点は人間がどう思うか。その統計量をどう使うか」

「同じようなことに悩んだ人たちがいると思うこと。その中の賢い人の考え方を使う」

「何を解決しなくてはならないかを自分で考え、それが見えてきたら、その方法を調べる」

「元の数字に戻れ、元のたくさんある数字を見て、自分の頭を使って何が気づきのようなものをさがせ」

「数字を扱うときは、データの処理のやり方をまず勉強するのではなく、自分のカンを活かし、何を知りたいかを考えて、それからそのやり方を調べる」

「数式を覚えるのでなく、数式の意味を直感でとらえる。人間が意味を理解すればあとは数式をコンピュータが覚えていてくれて、早く正確に計算してくれる。いくら計算できても、その数式にどういう意味があるのかわからなければその結果を使えない」

「標準偏差の出し方も昔は色々意見があったが、なるべく人間の直感にあうように試行錯誤して決めた」

「平均とはたまたまを取り払って、本当の姿を見ようとするもの」

「需要予測とはコンピュータが予測するのではなく、人間がカンを使って予測する。カンのない人は予測しちゃいけない」

「予測結果があたりそうか、あたりそうもないかは人間が判断する」

「カンを活かす道具が数学。すばらしいやり方を編み出した数学者もカンというひらめきがあり、それを数学で実証し、式をあみだした」

「ナレッジマネジメントとは試行錯誤して得た知識・ノウハウ・経験を誰でもわかる形に残していくこと」

「とびっきり頭が理論的な人が、他の人にぐうの音も出ないように理論化するのが数学」

「コンピュータは頭が悪いから、やり方を1通りにしないとできない。コンピュータ化されれば標準化される」

「微分なんてコンピュータがやってくれる。自動車がどうやって動くか知らなくても運転はできる」

「あたらないと困るなら予測なんてやるな」

「予測値に納得しないなら、予測のやり方を変えろ」

「予測のやり方は人間のカンがベースでそれを式にする」

「ビジネス数学は人間が持っている感覚、直感を数字で表すもの。強い・弱い、薄い・濃い、起きそう・起きそうもないといった感覚を数字で表す」

「予測精度の高い説明変数を選ぶのではなく、納得できる説明変数を選ぶ」

「データから考えて人間のイメージに合った式を選び、そのデータがたまってきた、状況がかわってきたら、ちがう式を使う」

「細かい計算などどうでも良くて、その考え方を理解する」

「あてるのなら直感の方があたる。今あるデータから考えるとこう考えられるはず」

「数字に表せないものなんてあるか」

「積分は面積を出すことで、コンピュータが全部やってくれる」

「考え方を理解しないで、自分が納得しないで数字を出したって、お客様も理解してくれないし、納得してくれない」

「めったに起きないことがたまたま起きるなんておかしい」

「コンピュータはただ数学者が考えた式どおりに計算しているだけだ。人間は頭を使って数学を使うんだ」

「もっと良いやり方はあるかもしれない。ただ今まで誰も気づいていない」