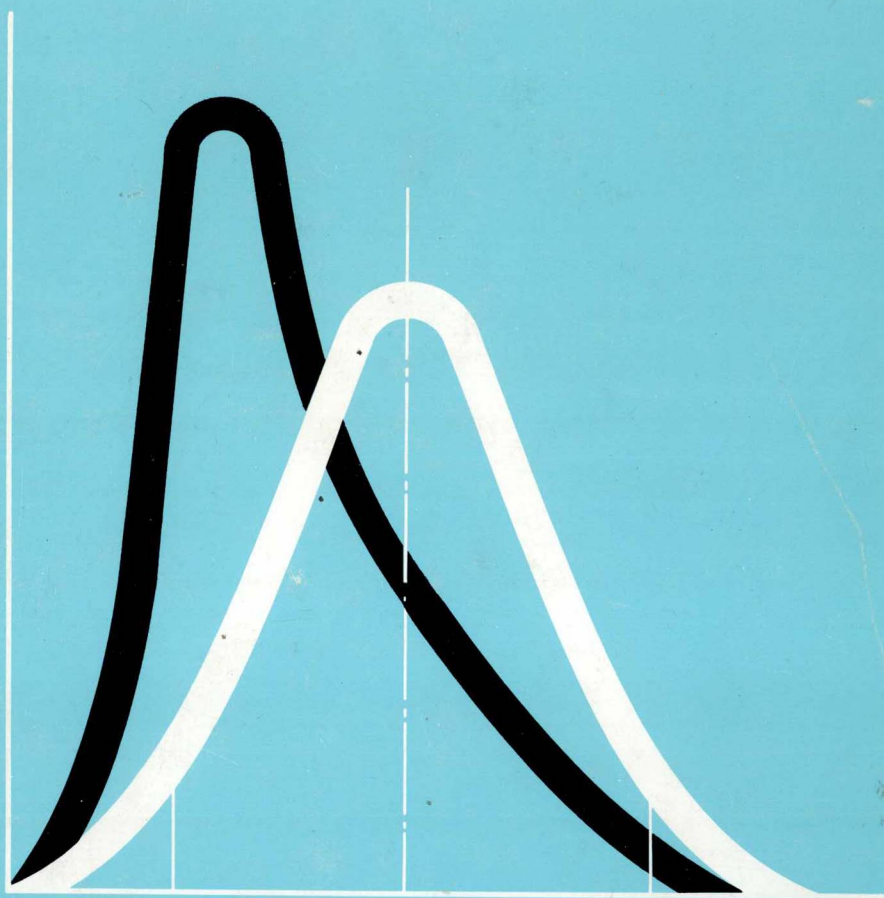


品質管理講座

新編

# 統計的方法

森口繁一編



品質管理講座

新編

# 統計的方法

森口繁一編

日本規格協会

## 森口繁一略歴

1916年 香川県小豆島に生まれる  
1938年 東京帝国大学工学部航空学科卒業，同講師  
1945年 同助教授  
1954年 工学博士  
1955年 東京大学教授  
1977年 東京大学名誉教授，電気通信大学工学部教授

## 主要著書・訳書

デミング著 推計学によるデータのまとめ方（1950年・岩波書店）  
数値表A, B（1954年，56年・日科技連出版社）  
品質管理（岩波全書，1955年・岩波書店）  
初等数理統計学（1957年・培風館）  
統計解析（岩波講座「現代応用数学」）（1958年・岩波書店）  
初等力学（1959年・培風館）  
JIS FORTRAN 入門上，下（1968年，69年・共立出版）

## 新編 統計的方法

定価 1,400円

1959年11月25日 初版第1刷発行  
1976年4月15日 新編第1刷発行  
1984年4月13日 新編第14刷発行

権利者との  
協定により  
検印省略

編者 森口繁一  
発行者 森五郎

発行所 財団法人 日本規格協会

〒107 東京都港区赤坂4丁目1-24  
振替 東京 6-195146 Tel 03 (583) 8001

印刷・製本 (株)第一印刷所 © Japanese Standards Association, 1959

ISBN4-542-50220-1 C3050 ¥1400E

## 新編へのまえがき

本書は日本規格協会「品質管理と標準化セミナー」の統計的方法のテキストである。

昭和48年11月にセミナーのカリキュラムが改正になり、実験計画法とくに直交配列を強化するなど統計的方法全体の時間数が増加したのにもなって、全面的に書き直しを行なったものである。例題は、旧テキストのものもかなり残っているが、時間の許すかぎり新しい例題でおきかえることに努めた。

本書は、暫定版の形で2年間セミナーで実際に使用し、その間の講師ならびに聴講生諸氏のご意見をとり入れて再改訂したものである。こうして、内容のいっそう充実した本書が従来にも増して広く統計的方法の普及に役だつことを念願する次第である。

昭和51年3月

森 口 繁 一

テキスト作成委員

---

監修 森口繁一（東京大学）

作成委員

小野藤男（日本オイルシール）

岸 暁男（武蔵工業大学）

君塚洋司（キリンビール）

工藤紀彦（東京芝浦電気）

古林 隆（埼玉大学）

清水龍蔵（第二精工舎）

芳賀敏郎（山陽国策パルプ）

広津千尋（東京大学）

川村正信（日本規格協会）

旧テキスト作成委員

---

監修 森口繁一（東京大学）

浦昭二（慶応大学）、市村裕（日東化学）、川村正信（日本規格協会）、喜多見 孟（早稲田高等学院）、戸田英雄（工業技術院）、芳賀敏郎（山陽国策パルプ）、三浦大亮（東レ）、吉沢正（山梨大学）

## 初版まえがき

本書は品質管理のための統計的方法を初めて学ぶ人を対象として作られた教科書である。

もともと、日本規格協会の 세미나 のテキストとして書かれ、回を重ねるごとに改訂されてきたものである。具体例を用いて手法とその意味を説明することに努めた点が特色であると思う。また、項目の選び方や扱い方は、「普通科」の教材としてどうしてもこれだけは必要だといえる程度にとどめ、よけいなことを入れないよう、また、配列も勉強しやすいようにできるだけ注意したつもりである。

私は石川馨氏とともに「統計的方法」担当の教務主任として、このテキストの企画・執筆指導・校閲にあたってきたのであるが、執筆は浦昭二、市村裕、芳賀敏郎、三浦大亮、などの諸君が、逐次にまたは分担して実行した。そのほか 세미나 講師として講義された多くの方々のご意見も随所に盛り込まれている。最後に活版刷の原稿の作成のためには喜多見孟君にとりまとめをやってもらった。それから統計量の分布に関する計算機による実験は戸田英雄、三浦大亮の両君の努力によるものであり、演習問題とその解答の作成は川村正信君の仕事である。実例の出所はそれぞれのところに注記したが、中には源泉のご希望に従って注記を控えたものもある。

以上に記したように、本書は大ぜいの協力の結果できあがったものである。ここにこれらの方々に感謝の意を表する次第である。

1959年11月15日

森 口 繁 一

## 目 次

|                                       |    |
|---------------------------------------|----|
| 1. 統計的な考え方 .....                      | 1  |
| 1.1 統計的な考え方                           | 1  |
| 1.2 データの性質                            | 3  |
| 1.3 統計的推測                             | 4  |
| 1.4 パラツキの原因                           | 7  |
| 演 習 問 題                               | 9  |
| 2. 度数分布 .....                         | 10 |
| 2.1 度数表                               | 10 |
| 2.2 層 別                               | 12 |
| 2.3 パレート図(曲線)                         | 13 |
| 2.4 欠点の生じやすい場所の調べ方                    | 14 |
| 2.5 区間分けによる度数表                        | 15 |
| 2.6 ヒストグラム                            | 17 |
| 2.7 平均と標準偏差の計算                        | 20 |
| 演 習 問 題                               | 22 |
| 3. 確率と分布 .....                        | 25 |
| 3.1 確 率                               | 25 |
| 3.2 確 率 分 布                           | 29 |
| 3.3 不良個数の確率分布(二項分布)                   | 37 |
| 3.4 欠点数の分布(ポアソン分布)                    | 39 |
| 3.5 正 規 分 布                           | 41 |
| 3.6 分散の加法性                            | 48 |
| 演 習 問 題                               | 51 |
| 4. 母平均に関する推測( $\sigma$ 既知) .....      | 54 |
| 4.1 サンプルの平均 $\bar{x}$ の分布             | 54 |
| 4.2 母平均 $\mu$ に関する仮説の検定( $\sigma$ 既知) | 66 |

|     |                                |     |
|-----|--------------------------------|-----|
| 4.3 | 母平均 $\mu$ の区間推定 ( $\sigma$ 既知) | 72  |
| 4.4 | 平均の差の推定と検定 ( $\sigma$ 既知)      | 75  |
|     | 演習問題                           | 79  |
| 5.  | 母分散に関する推測                      | 81  |
| 5.1 | 平方和                            | 81  |
| 5.2 | 不偏分散                           | 85  |
| 5.3 | 母分散の検定                         | 89  |
| 5.4 | 分散比の推定と検定                      | 92  |
| 5.5 | 範囲を用いる推定と検定                    | 98  |
|     | 演習問題                           | 103 |
| 6.  | 母平均に関する推測 ( $\sigma$ 未知)       | 105 |
| 6.1 | 母平均 $\mu$ の検定                  | 105 |
| 6.2 | 母平均の区間推定                       | 108 |
| 6.3 | 平均の差の検定 (対応のある場合)              | 109 |
| 6.4 | 二つの母平均の差の推定と検定                 | 112 |
|     | 演習問題                           | 118 |
| 7.  | サンプリング                         | 119 |
| 7.1 | 有限母集団からのサンプリング                 | 119 |
| 7.2 | 二段サンプリング                       | 122 |
| 7.3 | 層別サンプリング                       | 125 |
| 7.4 | その他のサンプリング                     | 129 |
|     | 演習問題                           | 131 |
| 8.  | 実験の計画と結果の解析                    | 133 |
| 8.1 | 一元配置 (母数模型)                    | 133 |
| 8.2 | 一元配置 (変量模型)                    | 141 |
| 8.3 | 二元配置 (繰返しなし)                   | 146 |
| 8.4 | 二元配置 (繰返しあり)                   | 154 |
| 8.5 | 分割実験                           | 163 |

|                    |     |
|--------------------|-----|
| 8.6 直交配列           | 169 |
| 演習問題               | 185 |
| 9. 2変数の間の関係        | 187 |
| 9.1 相関分析           | 187 |
| 9.2 散布図および2次元の度数分布 | 188 |
| 9.3 相関係数の計算        | 193 |
| 9.4 相関に関する検定と推定    | 198 |
| 9.5 回帰分析           | 201 |
| 9.6 回帰分析(繰返しのある場合) | 208 |
| 演習問題               | 214 |
| 10. 計数値に関する推測      | 216 |
| 10.1 不良率の推定        | 216 |
| 10.2 欠点数の推定        | 217 |
| 10.3 推計紙           | 218 |
| 10.4 符号検定          | 223 |
| 10.5 相関に関する推測      | 226 |
| 10.6 適合度の検定        | 231 |
| 10.7 分割表に関する検定     | 235 |
| 演習問題               | 242 |
| 付 表                | 245 |
| 演習問題解答             | 257 |

# 1. 統計的な考え方

この章では、初めて統計的方法を学ばれる人々のために統計的方法の基本的な考え方について述べ、さらにそれが品質管理に応用される際の一般的な基本原則にも触れる。

## 1.1 統計的な考え方

近代の品質管理の特徴として、二つの点があげられている。第1は統計的な考え方や手法が中心となって構成されていることであり、第2は組織的・全面的に遂行される点である。統計的品質管理ということばは第1の点を強調したものであり、アメリカの W. A. Shewhart の提唱した管理図法を始めとして、いろいろな統計的手法が縦横に駆使される。しかし品質管理は統計的な手法に偏して他の面への注意を怠ると、現場から浮いてしまい成果があがらない。工場全体が力をあわせて組織的に行うのでなければならない。

品質管理において統計的な考え方が重要なのは、品質というものがたえず変動するからである。どんな精密な機械でいくら注意深く作っても、かならず品物の品質にはバラツキ (dispersion) がある。しかし正常な状態にあればそのバラツキには統計的な規則性がある。品質の分布を統計的につかみ、統計的な推理にもとづいて行動するのが「統計的品質管理」の根本である。

統計的につかむというのは、個々のものに気をとられないで全体をながめるとのことである。そして一見不規則な変動の中から規則性をつかみ出すことである。ここで注意しなければならないのは、われわれが推論を下そうとしている対象は観測されたデータ (data) それ自身ではないことである。観測されたデータと、統計的な推理の対象となるものとは別である。

いま針金の受入検査の場合を考えてみよう。たくさんの針金の巻線の中からいくつかのサンプルを取り出し、それらの引張り強さを測定して、その結果から針金のロット(lot)全体を受け入れてよいかどうかを決定する。このとき測定された個々のデータが規格を満たしているかどうかには、当面の関心はない。それらのデータから見て、ロット全体が満足なものであるかどうかを推論したいのである。

注1. パラツキ：チラバリともいい、データの大きさがそろっていないこと。

パラツキの大きさを表わすには、たとえば標準偏差を用いる(20ページ参照)。

注2. ロット：製品・半成品または原材料などの単位体または単位量がある目的をもって集めたもの。

また別の例として、人絹糸の製造工程で毎日一定時間ごとに人絹糸の太さを測って、 $\bar{x}-R$  管理図を書いている場合を考える。人絹糸の太さを測るのは、その1本の繊維の太さを知るのが目的ではなく、こうして得られたデータを総合して、その製造工程の異常をできるだけ早期に発見しようとしているのである。

このようなことを強調するため、データをとってきたもとの集団——巻線のロットや人絹糸の製造工程——を母集団(population)と呼び、母集団からその一部を取り出したものをサンプル(試料, sample)という。統計的推理はサンプルをもとにして、母集団に関する推測をしようとするのである。

以上のことを整理すると、次の図1.1.1のようになる。

母集団についての推測は、行動の指針として現実の役に立つ。行動の決定

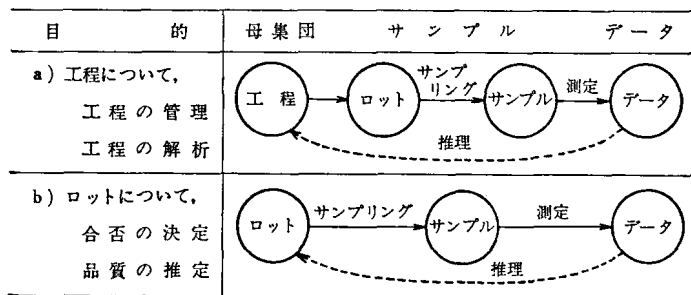


図 1.1.1 統計的推理

—— どのような行動をとるか、または、全く何の行動もとらないかの決定——には統計的推測の結果のみならず、固有技術の理論や経験、ならびに経済的その他の諸条件を考慮に入れた総合判断が必要である。

## 1.2 データの性質

統計的推理をするためにはデータがなければならない。データをとるのは何か目的があつてのことであり、統計的推理はその目的にそって行われるのである。ここでは統計的推理のもととなるデータの性質について考える。

われわれが得るデータは、計量値と計数値の二つに分けられる。計量値は連続量として測られる品質特性の値であり、計数値は不良品の数や欠点数などのように、個数で数えられる品質特性の値である。そのおのおのの例としては、次のようなものがある。

計量値の例：錠剤 1 個の重量(g)、針金の引張強さ(kg/cm<sup>2</sup>)、ガラス板の厚さ(mm)、電球の光束(ルーメン)

計数値の例：布 1 m<sup>2</sup> あたりのきずの個数、1 か月間の事故数、1 枚のガラス板の泡の数、塗装不良の部品数

なお丸棒の直径などは計量値であるが、通り止りゲージを用いて、ある一定幅におさまっているかどうかによって、これを良・不良に分けることもある。このようにすると、本来計量的なものが計数値(直径不良本数)として扱われることになる。計量値として測定するには時間や費用がかかるとき、簡易測定としてしばしばこの方法が使われる。測定 1 回当りの情報は少ないが、個数をふやせば同じ費用でより多くの情報が得られることがある。

データはかならずバラツキを持っているのであるが、その分布の中心を示すには通常平均値(mean)  $\bar{x}$  が用いられる。分布のひろがり、すなわちデータのバラツキの程度を定量的に示すには、範囲(range)  $R$  (最大値-最小値) が用いられる。たとえば、データが

35, 32, 34, 37, 38

と与えられたとき、平均値  $\bar{x}$  は

$$\bar{x} = \frac{35+32+34+37+38}{5} = 35.2$$

となり、範囲  $R$  は

$$R = 38 - 32 = 6$$

である。

ただし5個のデータから求めた  $\bar{x}$  と、100個のデータから求めた  $\bar{x}$  とが同じ値だといっても、それに対する信頼の度合は違ひし、また同じ母集団からとったとしても、たくさんのデータを観測すれば、それだけデータのひろがる範囲  $R$  は大きくなることが予想される。したがって  $\bar{x}$  や  $R$  を示すときには、かならずそれらが何個のデータから計算されたものであるかを示す必要がある。

注. データのパラッキを示すのに、標準偏差（不偏分散の平方根）

$$\sqrt{V} = \sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 / (n-1)}$$

を用いることがある（5.2節参照）。

データはそれをもとにして母集団に関する推理をするためのものであるから、できるだけよく母集団の姿を反映するようにサンプルをとらなければならない。特に小さな値とか、良品ばかりを取り出したのでは役に立たない。しかし完全に母集団の姿そのままをうつしだすようなサンプルを取り出すことは、たいていの場合不可能だし、もしそれができるならばサンプルを取って観測する必要もない。普通は母集団を構成する各部分がサンプルの中に含まれる機会が同じになるような抜取方法——すなわち**ランダム抜取** (random sampling)——を用いる。そうすると確率論にもとづいた客観的な推論を下すことができる。

### 1.3 統計的推測

母集団からサンプルを取って測定をする。そのデータから母集団に関する推測を行う。これが統計的方法の全般を通ずる原理である。サンプルは母集団の一部分であるから、このことは部分から全体を推測することに当たる。もちろん、部分を調べただけで全体が完全にわかるはずがない。たとえば、

サンプルの平均は母集団の平均から多少ずれた値になることはやむをえない。このずれが「サンプリング誤差」である。しかし、もしこの誤差が十分小さければ、それ相当の精度で部分から全体が推測できるわけである。ランダム抜取の手法による大規模な調査は、1930年代の後半のマハラノビスによるインド・ベンガル地方の作付の調査に始まる。ついで1940年代前半に、アメリカで大規模な社会調査がこの手法によって行われ大きい成果を収めた。戦後このようなサンプリング調査は、官庁統計はもとより、民間の各種統計調査・市場調査・世論調査の手法としても、標準的なものとして定着してきた。

例 1. 1944年アメリカで、急成長しつつある九つの都市の人口を、食糧・学校などの施策の基礎として、なるべく何か月もの遅れを伴わない方法で推計したいという必要に迫られた。そこで、たとえばロサンジェルスについては、全部で約2万個の街区(block)のなかからランダムに約3000個を選んで調査し、1街区当たりの人口の平均を推定し、これに街区数を掛けて全人口の推計を行った。

品質管理の面でも平均値ないし総量の推定のために、ランダム・サンプリングがよく用いられる。

例 2. 500 俵の羊毛の総重量を推定するために、ランダムに25俵を抜き取って目方を計った。その25俵の目方の平均が296kgと出たので、その500倍

$$296 \times 500 = 148000 \text{ (kg)}$$

をもって総重量の推定値とした。

例 3. コークスの灰分を知るため、1交替の間に3回、ベルト・コンベアの長さ1mの上ののっているコークスを採取し、これを縮分・分析してその1交替の間の平均灰分の推定値とする。

抜取検査においても、ロット全体のなかから一部分をサンプルとして抜いて検査し、その結果にもとづいて全体に対する処置——合格か不合格か——をきめる。このとき、サンプルの不良率がそのままロットの不良率ではないことに注意しなければならない。サンプルについて調べた結果にもとづいて

ロットの合格・不合格をきめる場合、あるいは不良率の低い良いロットが不合格になったり、あるいは不良率の高い悪いロットが合格になったりするおそれがある。このような危険が十分小さいように抜取個数や合格・不合格の判定の基準を決めるためにサンプルがランダムに取られるという前提のもとで、確率論を適用することができる。そのような理論によって抜取検査を設計し、かつその理論に合うように実際にランダム抜取を行って検査をし判定するようになったのは戦後のことである。

母集団とサンプルの関係で、最も注目すべき点は、推定の精度に主としてきいてくるのは抜取率よりも抜取個数であるという点である。たとえば、市場調査で、ある新製品に対する選好の度合を知るために、大きさ 5000 のサンプルを抜いて調べる場合、母集団の大きさが 100 万であっても、1000 万であっても、1 億であっても、サンプリング誤差はほとんど同じくらいである。抜取率でいって 0.5%、0.05%、0.005% といった違いがあっても、そしてそれがきわめて小さいにもかかわらず、サンプルの大きさ 5000 というものがきいてくるので、いずれの場合にもかなりよい精度が得られるのである。

また、たとえば、ある機械の稼働率を知るために、1 か月の間にランダムに選んだ 100 個ぐらいの時点で、その機械が動いていたか遊んでいたかを調べると、かなりの精度で稼働率のデータが得られるのであるが、この場合などは、抜取率という見方をしたのでは、無限にたくさんある時点のなかから 100 個の時点を選んだのだから、抜取率はほとんどゼロということになるわけであるが、それでもサンプルの大きさ 100 に相当するかなりの精度が確保できるのである。

工程の状況とか、ある操業条件のもとでの収率とかいったものも、相当長い時間にわたる平均的な性質を問題にすべきであろうが、上記と同じ理由で適当な時点で数百個のデータを取れば、実際上の行動の指針として十分な精度で推論ができるのである。

バラツキについても同様のことがいえる。すなわち、母集団の分布のひろがりぐあいを知るには、適当なサンプルについて、そのデータのバラツキの

程度を調べて推測することができる。そのときもサンプルの大きさがものをいう。

サンプリング誤差の大きさを意識しながら、母集団の性質について正当に推測できるだけのことをちゃんと推測し、その精度が許す以上の主張をしないこと。これが統計的推測にあたっての大切な心得である。

## 1.4 バラツキの原因

統計的品質管理の目的は品質のバラツキを小さくすることにあるということもできる。製品の品質に影響する因子は数限りなくある。その中で重要なものを見だし、それが管理できるものならできるだけ均一になるように努力し、管理できないものならそれに応じて他の因子を変化させて品質が一定になるようにする。

たとえば、人絹糸の太さはビスコースのセルローズ濃度、ビスコースの粘度など原料に関係するものや、紡糸機の吐出量など製造工程に関係するものによって決定される。それらの量が時間的、空間的に変動すると、それは人絹糸の太さのバラツキとなって出てくる。ある瞬間にほうぼうのノズルから出てくる人絹糸の太さは互に異なるし、また1本のノズルから出てくる糸の太さは時間的に変動している。したがって、いくつかのノズルから何回か人

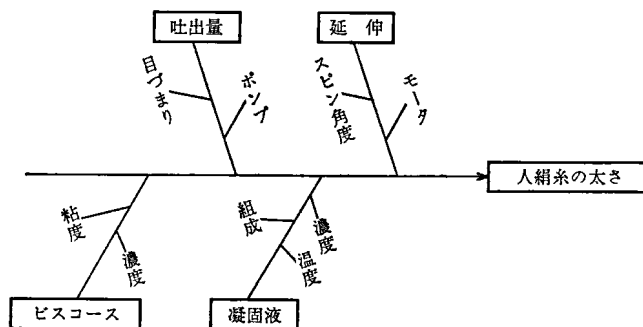


図 1.4.1 人絹糸の太さの特性要因図

網糸を取り出して太さを測定すると、そのような太さのパラツキを反映するデータが得られる。

無数にある変動原因をすべて管理することはもちろん望ましいことではあるが、それはとうてい不可能であり、またその必要もない。細かい変動はもしその原因を知っていたとしても、それを見逃してもかまわないが、大きな変動を与えるものはそれをとらえ出して管理しなければならない。ある品質特性について、工程の途中で変動原因がいくつも重なっていく関係を図示すると、たとえば図1.4.1のようになり、どのようなものから管理すべきかという目安を得るのに役立つであろう。図1.4.2に特性要因図の別の例を示した。

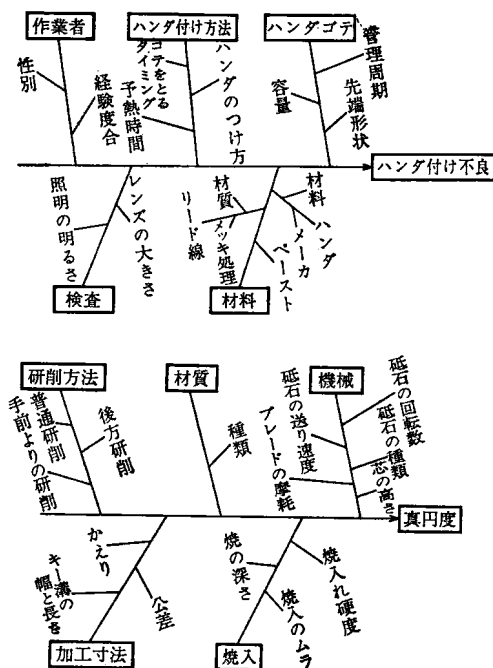


図 1.4.2 いろいろな特性要因図

演 習 問 題

- [1.1] 次の用語の意味を簡単に説明せよ。  
母集団, サンプル(試料), パラツキ
- [1.2] 計量値データと計数値データに分類せよ。
- (1) 溶液の濃度(%)
  - (2) 機械の故障回数(回)
  - (3) 真空管の寿命(時間)
  - (4) 錠剤の成分含有量(g)
  - (5) 合金板の引張強さ(kg/cm<sup>2</sup>)
  - (6) 鋼板の厚さ(mm)
  - (7) 電球の不良率(%)
  - (8) 鋼球の外径(mm)
  - (9) 塗装面のピンホールの個数(個)
- [1.3] 次のデータの平均値, 範囲を求めよ。
- (1) 5.5 5.0 5.1 5.6 5.2
  - (2) 55 50 51 56 52
  - (3) 10.095 10.090 10.091 10.096 10.092
  - (4) 1059.5 1059.0 1059.1 1059.6 1059.2