

## 図書館運営の評価と指標

岸田和明（駿河台大学 / 国立情報学研究所）  
kishida@surugadai.ac.jp

### 1. 評価の背景と目的

#### 1.1 図書館評価の目的

##### 図書館運営には評価が必要である

多くの図書館は非営利組織である。このため、一般の企業における売上高や経常利益のような業績評価のための単一かつ決定的な尺度はない。しかし、図書館においても、評価を導入することによって、その限られた経営資源（予算・職員・施設）の範囲内で効率的かつ効果的に業務を遂行し、よりよいサービスを提供していくよう努力する必要がある。

##### 計画策定と評価とは一体であることが望ましい

図書館評価には、

業務・サービスの計画策定（planning）中の一要素としての評価

現状や問題点を把握するための「探索的・発見的な」評価

の2種類がある。後述するように、評価には何らかの比較基準が必要である。この点、場合には、比較基準の設定が省略されることがあり、評価結果が十分活用されずに終わってしまうことが少なくない。実際の業務・サービスの改善を意図した明確な計画のもとで、評価を実施していくことが望ましい。

##### 自己点検・自己評価としての図書館評価

最近、多くの領域・場面で、自己点検・自己評価の必要性が認識され始め、実際に実践されている。

##### 行政評価（政策評価）としての図書館評価

公共図書館を国または地方自治体の行政・政策の一環として捉えれば、図書館評価を一種の行政評価として位置づけることが可能である。一般に、行政評価では、評価指標を

投入（インプット）          産出（アウトプット）          成果（アウトカム）

の3種類に分類する[1]。特に、最近、アウトカムの指標として、顧客満足度（Customer Satisfaction : CS）が注目されている（「満足度」とともに「重要度」も調査されることがある）。

#### 1.2 図書館評価の歴史的経緯

##### 図書館評価は長い歴史を持つ

図書館評価には長い歴史がある。日本においても、住民調査や来館者調査が、研究レベ

ル・実務レベルの両面において、頻繁におこなわれてきたし[2]、特に蔵書を評価するために書誌や貸出統計を利用する方法には、確立されたものがいくつかある。方法論的には、  
住民調査・来館者調査などのアンケート調査  
貸出統計を利用した数量的分析  
が代表的である。

### 米国における図書館評価の動向

図書館学の先進国である米国においては、早くから、一般市民の読書興味に関する調査や、図書館の平均的利用者像の調査が実施されてきた。また、貸出統計を利用した数量的分析の研究例も数多い。そのような伝統の中で、1980年代以降、各図書館が自ら評価を実施するための方法論の確立が模索されてきた。この背景には、例えば「住民1人あたりの蔵書冊数は最低3冊」のような具体的数値目標を、館種を限定したとしても、すべての図書館に共通に設定することは難しいという認識がある。つまり、規模・内容の点で多種多様な図書館に対して、それを取り巻くさまざまな状況を考慮せず、画一的な評価基準を設定することは困難とする考え方である。このため、

評価基準の画一化	評価方法の画一化
----------	----------

という一種の方向転換がなされた。この延長上に、図書館の評価指標に関する国際規格ISO11620（JIS X 0812が対応）がある。

### 図書館評価に関する国際規格

図書館評価に関する重要な国際規格は次の2つである。

ISO2789：図書館統計

ISO11620：図書館評価の指標

## 2. 評価の方法

### 2.1 目的と目標、達成目標

#### 目的と目標には階層構造がある

図書館評価を効果的におこなうためには、評価対象となる業務・サービスの目的・目標の明確化が欠かせない。目的・目標は図1に示したような階層構造を持っている[3]。公共図書館ならば、その最上位に図書館法第2条「……図書、記録その他必要な資料を収集し、整理し、保存して、一般公衆の利用に供し、その教養、調査研究、レクリエーション等に資することを目的とする……」が位置することになる。最上位の目的を**使命**（mission）と呼ぶ。

#### 階層構造の持つ意味

使命は一般に抽象的である。このため、そのままでは役に立たず、使命から具体的な方針を導く必要がある。この導出をわかりやすく、全体として整合的におこなうためには、いくつかのレベルで目的・目標を設定して、それらを階層的に展開していく必要がある。当然、各目的・目標はその上位の目的・目標によって制限され、上位の目的・目標に抵触するような下位の目的・目標を設定することはできない。その最下層はしばしば**達成目標**と呼ばれる。

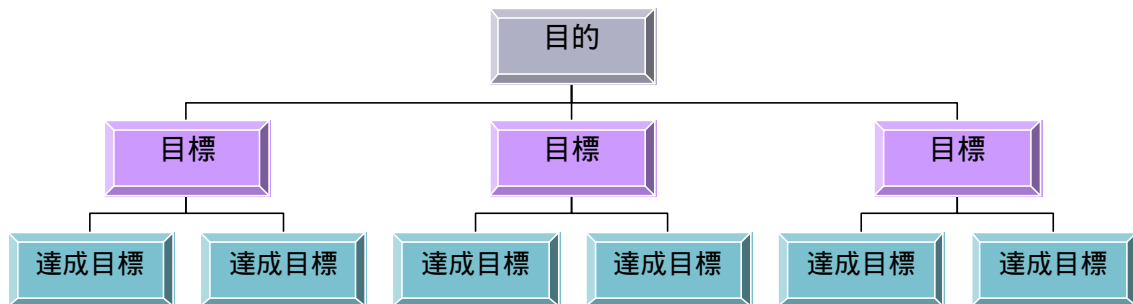


図1 目的・目標・達成目標

### 達成目標には具体的な数値を設定することが望ましい

達成目標を評価基準として使うことができる。しかも、もし、「住民1人あたりの貸出延べ冊数を3冊にする」「1日当たり20件の目録データを作成する」といった数値を達成目標に盛り込むことができれば、評価作業は容易になる。

## 2.2 効果と効率、パフォーマンス

### 「効果」と「効率」は異なる観点からの評価である

目的や目標が実際に達成された程度を**効果** (effectiveness) と呼ぶ。例えば「1日に作成する目録データの件数向上」を目標とした場合に、

システムA：1日30件    システムB：1日20件

ならば、システムAのほうが効果的である。一方、目標や目的の達成に要した資源量に関する**効率** (efficiency) も考慮しなければならない。例えば、システムCとシステムDはともに1日30件作成できるが、その経費が、

システムC：1日5,000円    システムD：1日7,000円

ならば、システムCのほうが効率的である。このように、

- ┌ 効果：目的や目標が達成された程度
- └ 効率：目標や目標の達成に要した資源の量

であり、これらは異なる側面を評価するものである。なお、この効果・効率を総称して、**パフォーマンス** (performance) と呼ぶことがある。

### 費用対効果の概念とは

ところで、上の例で、

システムA：1日6,000円    システムB：2,000円

とすれば、効果の観点ではシステムA、効率の観点ではシステムBが優れていることになる。この場合にデータ1件あたりの経費を計算すれば、

システムA：1件あたり200円    システムB：1件あたり100円

となり、システムBが優る。つまり、**費用対効果** (cost-effectiveness) の観点からは、システムBのほうが優れていることになる。このように費用対効果とは、いくつかの代替案（この場合にはAとB）があったときに、その効果だけではなく、費用をも含めて判断するという考え方である。なお、もちろん、1件あたり100円程度の差ならば、一日に30件作成できるAを選ぶという選択もできる。

## 2.3 測定と尺度、指標

### 評価対象の概念的定義と操作的定義

例えば「図書館資料の利用を促進する」という目標を設定し、実際に「資料の利用」が増えたかどうかを評価したいとする。しかし、「図書館資料の利用」は抽象的（概念的）すぎて、このままでは、その量を評価することはできない。そこで「図書館資料の利用」を「図書館に来館した人がその図書館資料を手にとって目を通した場合」のように具体的に定義する必要がある。このような定義を**操作的定義**（operational definition）と呼ぶ。

### 尺度の設定と測定

操作的定義を設定したら、次に**尺度**を考えなければならない。この場合には、「図書館資料の利用の促進」が目標であるから、例えば、

「利用回数」...資料が利用された場合の数

「利用者数」...資料を利用した人の人数

などが尺度の候補となる（ここでの例とはややずれるが、図書館資料を利用した結果としての「満足度」、その資料が実際に役立ったという意味での「有用性」などの尺度も考えられる）。尺度が設定されれば、それを実際に**測定**する。例えば、図書館に来館した人に調査票を配布し、その日利用した資料の数について回答してもらえば、利用回数（または利用者数）を測定できる。

### 評価に伴う誤差には標本誤差と測定誤差とがある

しかし、(1)すべての来館者に対して調査することは難しく、また、(2)来館者の記憶違いなどの誤りが発生することがある。(1)は全体に対して調査できず、部分的な標本に対する測定にならざるを得ないことに起因する誤差であり、一般に、これを**標本誤差**（sampling error）と呼ぶ。一方、後者(2)は測定の際に混入する誤差であり、**測定誤差**（measurement error）と呼ばれる。

### 尺度の妥当性と信頼性に注意しなければならない

来館者への調査には測定誤差が含まれ、再度質問した場合にその回答が変わる可能性がある。この点で、この回答から集計される尺度の**信頼性**は低い。一方、尺度として貸出回数を使ったとする。貸出処理の誤り率は非常に低いという仮定の下で、この尺度の信頼性は高い。しかも、すべての貸出業務を対象に貸出回数を集計すれば、標本誤差も生じない。しかし、「図書館に来館した人がその図書館資料を手にとって目を通した場合」という操作的定義と、「貸出」の定義とには、若干のずれがある（目を通したからといって借り出すとは限らない）。この点で、貸出回数の**妥当性**は低い。妥当性・信頼性がともに十分な尺度を使えばそれに越したことはないが、ここでの例のように妥当性と信頼性のトレードオフが存在するような場合があり、尺度の選択に注意を要する。

### 測定結果から指標を計算する

測定結果が得られたならば、評価したい事項に合わせて、何らかの**指標**（indicator）を設定して計算する。例えば、「蔵書1冊あたりの貸出回数」は、蔵書規模の異なる主題分野間での利用の程度を比較するための指標である（「蔵書冊数」を蔵書規模の尺度、「貸出冊数」を利用程度の尺度として、後者を前者で割ったものがこの指標であり、これはしばしば蔵書回転率と呼ばれる）。

## 名義尺度・順序尺度・間隔尺度・比尺度

一般に、その特徴に従って、尺度を次の4種類に分類することができる。

名義尺度：例 - 「男」「女」

順序尺度：例 - 「満足である」「ふつう」「不満である」

間隔尺度：数値であるが、原点を持たず、割り算に意味のないもの：例 - 気温

比尺度：原点を持つ数値で、割り算に意味のあるもの：例 - 体重

アンケート調査の場合に、順序尺度に数値を割り当て(「満足」:3、「ふつう」:2、「不満」:1)、平均値等を計算することもある。

## 2.4 評価の手順

### 評価自体の目的を明確にする

実際の評価の手順は図2のようになる[3]。最初に、評価自体の目的をしっかりと決めておくことが重要である。評価の目的や対象が何であるのか、評価結果を何のために使うのかをあらかじめ明確にしておかないと、評価作業を進めていくうちに、いったい何を何の目的で評価しているのかがわからなくなるといことになりかねない。

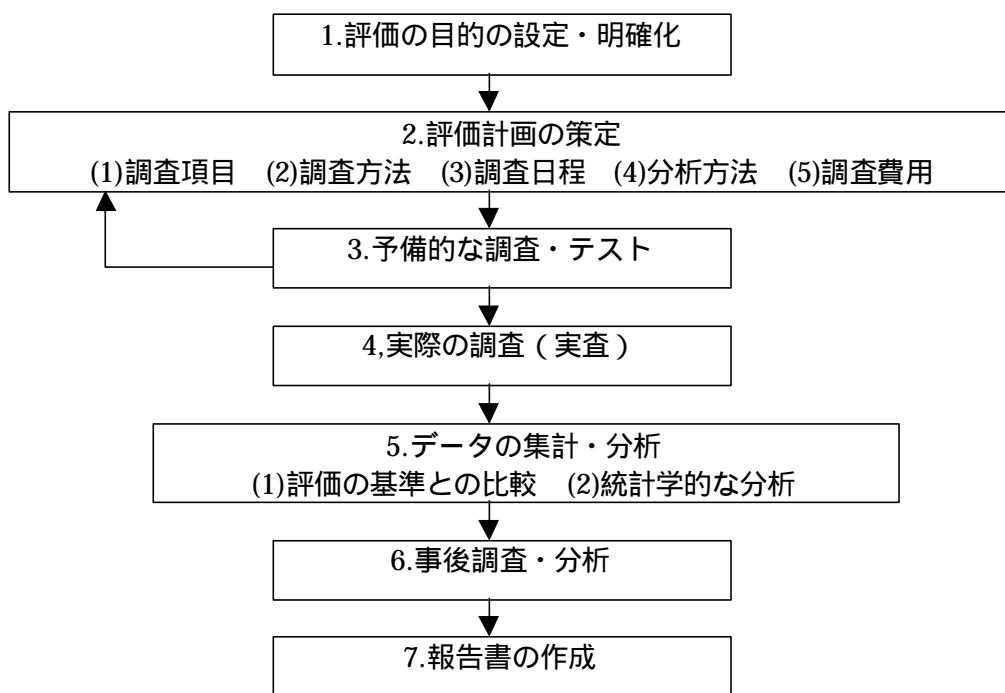


図2 図書館評価の手順

### 2 種類のデータ - 「業務統計」と「調査統計」

評価のためのデータ収集方法としては、

業務の遂行において生成される記録の集計

アンケート調査などの特別な調査の実施

の2つがある。前者の方法で集計されるデータを業務統計と呼び、後者のデータを調査統計と呼ぶ。図書館における業務統計の例は貸出統計であり、調査統計の例は来館者に対する満足度のデータである。業務統計は調査のための特別な経費がそれほどかからないのに対して、調査統計はかなりの経費が必要になる。一方、調査統計が評価したい事項を

直接調べられるのに対して、業務統計として集計できる事項はかなり限定されている。

### 予備調査の重要性

評価のために住民調査や来館者調査などのアンケート形式の調査をおこなう場合、あるいは、貸出統計のような業務統計を集計して評価をおこなう場合のいずれにおいても、予備的な調査が欠かせない。特に、アンケート形式の場合には、調査票の設計に予備調査が重要な役割を果たすことが少なくない。

### 評価の基準やガイドラインとの比較

図3中の5.(1)「評価の基準との比較」に関しては、例えば、  
業務・サービスの具体的な達成目標と比較する  
異なる業務・サービス間で比較する  
規模や性質が類似した他の図書館と比較する  
IFLA（国際図書館連盟）などによる外的な基準と比較する  
などの方法がある[3]。このうち に関しては、日本においては、『日本の図書館』や『日本図書館年鑑』（共に日本図書館協会発行）に公共図書館や大学図書館の各種統計が掲載されているので、それらを用いることができる。また、 に関連して、日本における最近の例としては「公共図書館の設置及び運営上の望ましい基準について」の参考資料中に数値目標が掲載されている（その他の基準については『図書館法規基準総覧』（日本図書館協会）を参照）。

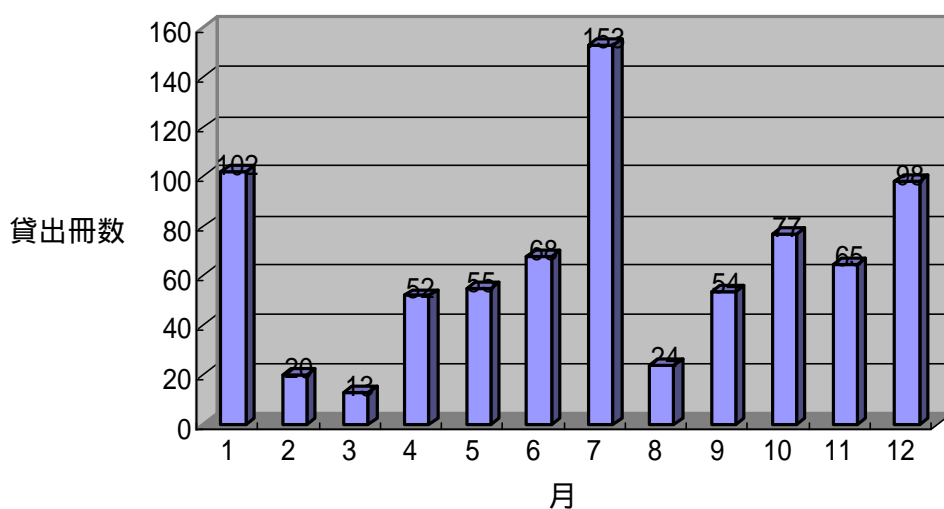


図3 月別での貸出延べ冊数（例）

## 2.5 記述統計と推測統計

### 図や表を作成して眺めてみる

評価のために収集したデータは、経営計画に示された達成目標や各種の基準・ガイドラインと単に比較するだけでなく、さまざまな角度から探索的・発見的に分析することが可能である[3]。その第一歩として、分布（distribution）に関する図や表を作成して、眺めてみるのが重要である。例えば、図3は貸出延べ冊数を月別に集計して、グラフ化したも

のであり、このようなグラフ（あるいは表）から何らかの規則的な傾向を発見できるかもしれない。

#### マクロな統計量を利用する

図やグラフではなく、平均値や中央値（メジアン）のような統計量を計算したほうが便利な場合もある。このようなマクロな統計量を有効活用すれば、複雑な状況をコンパクトにまとめて分析できる。ただし、その見返りとして、マクロな統計量は、本来の分布の持つ情報を何らかのかたちで捨ててしまっている点に注意しなければならない。

#### 標本誤差の計算が必要となる場合がある

データの収集方法として標本調査が採用されている場合には、標本誤差を考慮するために、推定・検定などの推測統計学の理論・手法を適用しなければならない。業務統計の分析において、その業務の範囲内の全数調査とすることが可能ならば、推定・検定を伴わない記述統計学で十分である。一方、住民調査等においては、多くの場合、標本誤差の推定が必要になる。この場合には、統計学の専門家の支援を受けることが望ましい。

## 2.6 質的な評価

### 質的な評価の方法

数量的な尺度や指標を用いる以外に、利用者や専門家、あるいは職員の意見に基づいて評価をおこなうこともできる。よく使われる方法は、アンケートの中に「ご意見がありましたらご自由にお書きください」のような、自由回答形式の質問項目を設定することである。これによって、調査表で設定された調査項目の枠にとらわれない意見をきくことができる[3]。

## 3. 評価のための統計と指標

### 3.1 業務統計の種類

#### 代表的な業務統計

すでに述べたように、日常的な業務上の記録（トランザクション）から集計される統計を業務統計と呼ぶ。図書館における業務統計は、大きく、次の3つに分けることができる。

- 所蔵資料に関するもの
- 提供しているサービスに関するもの
- 資源に関するもの

それぞれの代表的な統計を表1に示す[3]。

#### 業務統計を必要に応じて細分して使う

主題分野別あるいは利用者別などの区分に従って業務統計を細分して集計すると、評価や現状分析に有効な場合が少なくない。このためには、適切な区分を設定して、その情報が集計の源となる業務記録に含まれていることが必要である（例えば、主題分野別の貸出延べ冊数を集計するには、貸出トランザクションにNDCを何らかの方法で組み込む必要がある）。

表 1 代表的な業務統計

種別	業務統計
資料	蔵書冊数、年間受入冊数、年間除籍冊数、雑誌購入種数
サービス	開館日数、入館者数、登録者数、貸出冊数、相互貸借件数（貸出冊数、借出冊数、文献複写受付件数、文献複写依頼件数）、電子複写枚数、参考調査業務の受付件数・回答件数
資源	専任職員数、非専任職員数、経常的経費、人件費、資料費、図書購入費、製本費

### 3.2 主要な評価指標

#### 伝統的な評価指標の例

業務統計を利用した伝統的な評価指標を表 2 に示す[3]。

表 2 業務統計に基づく伝統的な評価指標

指標	計算方法（定義）
蔵書新鮮度	受入冊数 ÷ 蔵書冊数
貸出密度	貸出延べ冊数 ÷ 定住人口（= 実質貸出密度 × 登録率）
実質貸出密度	貸出延べ冊数 ÷ 登録者数
蔵書回転率	貸出延べ冊数 ÷ 蔵書冊数

#### 蔵書回転率の使い方

例として、蔵書回転率の使い方を示す。例えば、ある図書館において、分野 A と分野 B の貸出延べ冊数（1 年間）が、

分野 A : 100 冊 分野 B : 200 冊

であったとする。この結果から直ちに分野 B の貸出が相対的に活発であるとは結論できない。もし、それぞれの蔵書冊数が

分野 A : 50 冊 分野 B : 200 冊

であれば、分野 B の貸出延べ冊数の多さは単に蔵書規模の大きさに起因する可能性がある。このことは、蔵書回転率を算出すれば、

分野 A : 2.0 冊 分野 B : 1.0 冊

であり、一層明確になる。さらにもし、蔵書回転率の低い分野の相互貸借が活発ならば、選書に問題のある可能性が強く、一方、相互貸借もおこなわれていなければ、その分野が本質的に不活性である（少なくともその図書館の利用者にとっては）と判断できる。

#### 国際的な標準規格における評価指標

図書館の評価指標に関する国際的な標準規格である ISO11620 (JIS X 0812) において規定されている主な指標を表 3 に示す。この表中で、「タイトル」は、ある一冊の図書やある一冊の雑誌などの「ある一点の資料」を意味する。

#### 評価指標を計算するための特別な調査

表 3 には「利用者満足度」などの業務統計としては得られないデータから算出する指標がいくつか含まれている。このような評価指標を計算するには、次のような調査が必要になる[3]。

業務記録に若干の項目を付加する程度の調査



利用者の協力を必要とせずに、職員が図書館内で実施可能な調査  
 利用者の協力を得て、図書館内で実施する調査  
 図書館外の外的な事象や要因を対象とする調査

表3 ISO11620で規定されている主な指標

サービス、活動、あるいは その他測定されるもの	指標
(a)利用者の意識	
(a-1)全般	利用者満足度
(b)利用者サービス業務	
(b-1) 全般	特定サービス対象者の利用率，利用者当たり費用，人口当たり来館回数，来館当たり費用
(b-2) 資料の提供	タイトル利用可能性，要求タイトル利用可能性，要求タイトル所蔵率，要求タイトル一定期間内利用可能性，人口当たり館内利用数，資料利用率
(b-3) 資料の出納	閉架書庫からの資料出納所要時間（中央値），開架からの資料探索所要時間（中央値）
(b-4) 資料の貸出	蔵書回転率，人口当たり貸出数，人口あたり貸出中資料数，貸出当たり費用，職員当たり貸出数
(b-5) 他の図書館からの資料提供	図書館間貸出の迅速性
(b-6) レファレンスサービス	正答率
(b-7) 情報検索	タイトル目録探索成功率，主題目録探索成功率
(b-8) 設備	設備利用可能性，設備利用率，座席占有率，コンピュータシステム利用可能性
(c)整理業務	
(c-1) 資料の受入	受入に要する時間（中央値）
(c-2) 資料の整理	整理に要する時間（中央値）
(c-3) 目録業務	タイトル当たり目録費用

### 利用可能性の調査方法

利用可能性（availability）は、ある特定のタイトルが実際にその図書館で入手可能かどうかを意味する概念であり、長年にわたって、研究・議論されてきた。この利用可能性を測定するための最も簡単な方法は、

何らかの書誌を選び、

その書誌中の文献がその図書館に実際に所蔵されているか、あるいは、それが書架上に存在するかを調べる、

ことである。この方法ならば図書館員だけですべて実施することができる。一方、実際の利用者に協力を依頼し、利用者が実際に要求した文献に対して、上記を調べる方法もある。さらに、その文献が所蔵されていない場合に、相互貸借などの方法で後日利用可能になることもあるので、例えば7日とか10日とか期間を限定し、その期間内の利用可能性を調査することもできる。

### 3.3 来館者調査と住民調査

### 図書館に実際に訪れた人に対する調査：来館者調査

来館者調査の場合、実際に図書館に訪れた人に対して、利用したサービスやそれに対する満足度などを調べる。図書館の入口で調査票を配布し、利用者が帰る際にそれを回収すればよく、後述する住民調査よりも実施が容易である（標本抽出の必要はなく、また、回収率も住民調査に比べれば高い。経費も比較的かからない）。反面、非利用者に対する調査はできない。また、特定の曜日や時間帯にのみ実施した場合、回答が偏る可能性がある。

### 一般の住民に対する調査：住民調査

住民調査の場合には、一般に、標本を抽出し、調査票を郵送する必要がある。調査が複雑かつ経費がかかる反面、図書館に実際に足を運ばない人々の意見や考え方を知ることができる。

### 調査票の設計における注意事項

来館者調査や住民調査には、調査票の設計が重要である。調査票の設計における基本的注意事項は以下のとおりである[4]。

- 不必要な質問は含めないこと

- 質問を設定する前に、それからどのような統計ができるかを想像してみること

- 質問あるいは説明の言葉使いは平易な話し言葉そのままに

- 回答が難しい表現を必要とするものは避けること

- 質問によって回答を誘導することがある

- 複数の内容をもった質問は避けること

- 個々の質問の無回答を極力なくすよう努力すること

- 質問の量は多すぎないこと

### 質問の形式

例えば、「あなたは 〇〇 に満足していますか」と質問し、「たいへん満足している」「満足している」「あまり満足していない」「満足していない」といった多段階で回答してもらう質問形式が一般的である。これは順序尺度であり、それぞれの回答比率を計算すれば、傾向を把握できる（適当に重み付けして平均を計算することもある）。また、期待していたサービスの水準と実際に受け取ったサービスの水準との格差（ギャップ）を調査するような複雑な調査方法もある[1]。

### 参考文献

[1]系賀雅児「図書館評価をめぐる諸概念の相互関係」『第50回日本図書館情報学会研究大会発表要綱』日本図書館情報学会，p.43-46，2002．

[2]森耕一編『図書館サービスの測定と評価』日本図書館協会，1985．

[3]岸田和明「第9章 図書館業務・サービスの評価」『図書館経営論』改訂版，高山正也編，樹村房，2002．

[4]浅井晃『調査の技術』日科技連，1987．