歳時記



## あけび

果肉は甘いので食用にし，若葉はゆでたり，塩づけにして食べられる。木部にはアケビンが含まれており，生薬名を〈木通（もくつう）〉といって，煎じて利尿，通経剤として利用されている。

ところで，アケビの語源には色々の説があるようだ。代表的な説としては，果実が熟すると縦に裂けて白い肉をあらわすことから開け実，口を開けた様子がちょうど欠伸（あくび）や，さらに成熟した女性の中枢，開けつびを連想することから，開け実・欠伸（あくび）・開けつびがアケビとなったという説である。

味の方はどうかというと，これが何とも言えない。見掛けよりはズーッとおいしそうです。でもアケビの味は食べてみなければわかりません。

## 今月のおもな行事

| | |
|---|---|
| ○1日 | 漁業センサス調査日 |
| ○3日 | 文化の日 |
| ○6日 | 茨城県統計大会（県民文化センター） |
| ○9～10日 | 漁業センサス審査集計ブロック会議（大子町） |
| ○23日 | 勤労感謝の日 |
| ○28～29日 | 消費動向調査ブロック会議（静岡県） |
| ○29～30日 | 漁業センサスブロック会議（東京都） |
| ○30日 | 法人企業投資動向調査日 |

さいじき

# 標　準　偏　差

## §1．共通1次試験と標準偏差

世の中のしくみには，メスをいれて大胆な手術をしなければならない面が少くないようですが，そうかといってやたらにメスをいれるのも考えものです。塗り薬か何かですませられれば，それにこしたことはありません。さしずめ来年度から実施されることになった国立大学入試における共通学力第1次試験などは，傷口をよく消毒して，塗り薬を塗りこもうといったところでしょう。この種の治療においては，性急に効果を期待するのは困難です。しんぼう強く気長に療養することが肝心です。――とはいうものの，現に入試を控えて猛勉強中の受験生諸君やその周囲の人たちは，さまざまな期待と不安のもとに54年1月13日，14日実施ときまった共通1次試験に備えられていることでしょう。

来年度からの国立大入試については，当事者はもとより，一般の方々もすでによくご存知のように，この共通1次の成績が大きくものをいうことになります。その扱いは大学によってまちまちで，一般大学ではさらに大学独自の入試(第2次試験とよぶことにします)を行い，1次と2次の結果を総合して合否をきめることになりそうです。いずれにしても，共通1次が重要な意味をもってきますので，大学入試センターでは来春の本番に備えて，試行テストを行ってきました。その大がかりなテストが昨年行われ，今年の5月末にその結果を公表しています。それによると，テストの得点分布図はたいへんきれいな正規曲線を描いているということです。大体，この種の標準的なテストでは，多人数の得点分布は正規分布をするのがふつうで，そうでないと標準テストとは認め難いといわれています。

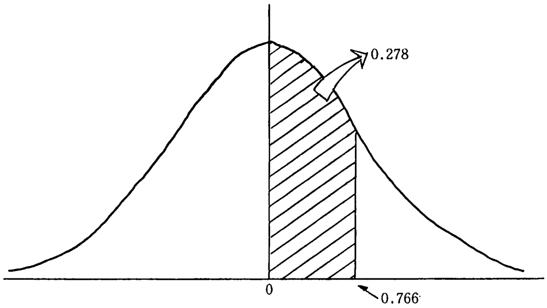
ところで正規分布は，平均と標準偏差でピシッときまる分布です。たとえばテストの総合点が平均555点であるからといっても，650点をとった人は，上からどの程度のところに位置するかはわかりません。平均以上であることは間違いないのですが，もしこのテストの標準偏差が124であったということがわかれば，650は標準偏差を尺度にして，平均よりも

$$\frac{650-555}{124}=0.766$$

だけ右にあることがわかります。それで統計の本などの巻末に載っている正規分布表をつかって所要の確率を読みとることができます。すなわち図―1の斜線の部分の面積が数表から約0.278と読めますから，650点以上の人が全体の

$$0.5-0.278=0.222$$

だけいることになります。したがってこの人は，上位22%ぐらいのところに位置していることがわかります。




図―1　正規分布表からの読みとり

先般公表された共通1次試行テストの結果は，総合点が平均555.69，標準偏差124の正規分布にしたがうということでした。従って受験者40万人，合格者8万人を想定し，上位20%以内にはいることを考えますと，

$$555.69+124\times0.842=660.10$$

以上の得点が必要ですし，受験者が60万人の場合には上位13.33%が合格というわけですから，

$$555.69+124\times1.111=693.45$$

以上の得点が必要になる，というようなことが新聞などで報道されていました。上で8万人という数を使っていますが，これはもちろん国立大学の入学定員です。また0.842とか1.111は正規分布表から読みとった値です。たとえば0.842についてですが，上位20%というのは正規分布表の斜線部分の面積が

$$0.5-0.2=0.3$$

ということですから，その横軸の値を表から(逆に)読んで，0.842を得るのです。このように，受験者の平均と標準偏差がわかれば，各自が自分は全体のどの程度のところにいるかがわかるわけです。しかし本番では平均点は公表されますが，標準偏差は公表されないようです。それでは上のような計算ができないではないかという不安が残りますが，心配はいりません。平均はテストいかんによってかなり変

東京理科大学教授　牧　野　都　治

動するかもしれませんが，標準偏差の124という値はあまり大きく変わらないと思ってさしつかえないからです。(注)原稿作成時は，まだ標準偏差の公表は決められていなかった)

ただし前に述べたことは，すべて共通1次だけについてのことであり，また志願する大学によって難易の違いがありますので，これは単なる目安でしかないことはいうまでもありません。

それでは第2次試験はどうなるのか。これにちなんでC大学での笑いばなしをひとつご紹介してみましょう。

C大工学部でも，かねがね入試制度を何とかしなければという考えはあったのですが，それでは大改革をめざして，学部をあげて取り組むかというと，そんな意欲はないようでした。むしろ国立大入試の共通1次が実施されるという方向に固まってきた時点あたりから，国立大協会へのアンケートに答えたり，第2次入試をどうするかを検討したりしなければならないという必要に迫られて，急きょ入試委員会を設置したという有様です。しかし，これはあながちC大に限らず，どこの大学も似たり寄ったりといったことのようでした。

さて，その委員会でのできごとですが，第2次試験は数学だけにしようということになりました。それでは，その配点をどうするかこれがなかなかまとまりません。

「1次でも数学があって，1,000点満点中200点なのだから，2次もせいぜい200点が限度だろう」

と主張して譲らないのは，何と出題側の数学の先生たちです。これに対し，他の先生方は

「1次1,000点，2次200点というような配点では，1次の結果でほとんど決まってしまい，わざわざ2次試験をやる意味がなくなりはしないか」

と心配しているのです。

そのとき，委員長がつぶやくように言われました。

「第2次試験で逆転可能かどうかは，まさに第2次での成績のバラツキがどうなるかということとかかわりがあるんではないですか。それで参考までに，いまお配りした試行テストの結果を見ているのですが，それによると数学は200点満点で平均112.72，標準偏差45.11とありますね。われわれはどうも200点満点ではピンとこないのですが，100点に換算すると標準偏差はいくらになると考えたらよいのですか。」

すると，ふだん実験データの解析などで統計を縦横に駆使していると言っておられるA教授の講座の助手のBさんが，

「200点で45.11ですから，100点満点ではその半分の……」

と言いかけたのにむかってA教授，

「バカだなぁ。君はすぐそういうオッチョコチョイなことをいう」と，あわてて制止して

「100点満点のXと100点満点のYを加えると，分散はそれぞれの和，つまり200点満点では分散が100点満点のときの2倍になるんだから，逆に200点を100点に直すと，分散が$\frac{1}{2}$になります。ですから標準偏差は$\frac{1}{\sqrt{2}}$倍，つまり$45.11\times\frac{1}{\sqrt{2}}$と計算して，大体32ぐらいですね」

と解説された。委員長はびっくりした様子で，

「標準テストにしては，意外にバラツキが大きいものですね。ふつうの入試でさえも，そんなにはバラツカないと思うんですが……。まあ，そんなことをふまえて，第2次試験の重みをどうしたらよいかを，次回あらためてご検討いただくことにしましょう」

と，この日はそれでチョンということになりました。

しかし会議終了後，A教授のもとに

「あれは，やはりBさんの説明が正しいようですが……」

と若い先生から電話があり，A教授もハッとして，思わず頭をかいたということです。

このあやまりはどこにあったかといいますと，しごく簡単なことでした。それは，

『確率変数Xの分散をV(X)とすると，CXの分散は

$$V(CX)=C^2\cdot V(X)$$

したがってCXの標準偏差はXの標準偏差のC倍になる』

という考えで計算すべきところを，勘違いして

「XとYとが独立であれば

$$V(X+Y)=V(X)+V(Y)$$

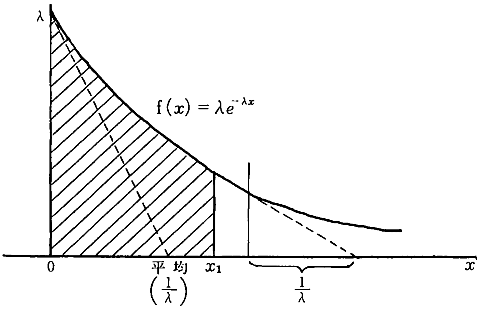
が成り立つ」

という定理(これも正しい定理ですが)を用いようとしたことによるものです。

さてその後の委員会で，第2次試験の配点がどう決められたかは，つまびらかでありません。しかし，共通1次での(1,000点満点で)標準偏差が僅かに124程度ですので，それ程つよい重みを与えなくても，第2次試験で十分逆転可能にはなるようです。

## §2．記憶喪失型の標準偏差

前節では正規分布を背景に話を繰り広げてきたのですが，こんどは正規分布と異なり，左右が非対称の分布について考えてみましょう。それらの中で，とくに重要なのは指数分布です。




図―2　指数分布の性質（1）

これは図―2のような密度曲線を持つ分布であって，確率変数Xの密度関数$f(x)$が

$$f(x)=\begin{cases}\lambda e^{-\lambda x} & (x\geqq 0)\\ 0 & (x<0)\end{cases}$$

で表わされます。この分布の平均は

$$E(X)=\int_0^\infty x\cdot\lambda e^{-\lambda x}dx=\frac{1}{\lambda}$$

また

$$E(X^2)=\int_0^\infty x^2\cdot\lambda e^{-\lambda x}dx=\frac{2}{\lambda^2}$$

ですから，分散は

$$V(X)=E(X^2)-E^2(X)=\frac{1}{\lambda^2}$$

したがって標準偏差は平均と同じ$\frac{1}{\lambda}$になります。

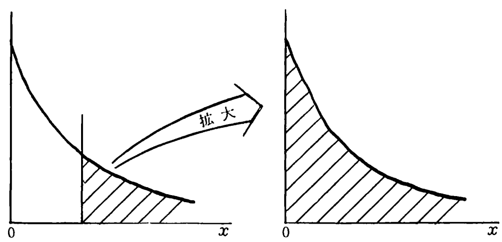
$$\frac{\text{標準偏差}}{\text{平　均}}=\text{変動係数}$$

ですから，指数分布の変動係数はつねに1であるということにもなります。

ところで，図―2で$x=0$の点での曲線への接線をひいてみますと，それが横軸とちょうど$\frac{1}{\lambda}$のところで交わります。平均も標準偏差も$\frac{1}{\lambda}$なのですから，図―2が正しく書けていると，平均も標準偏差も図の上からすぐにけんとうがつくという，おもしろい性質をもっています。しかし指数分布の何よりもきわだった特徴は，それが「過去の履歴に無関係な」いわゆる〝記憶喪失型の分布〟であるといってよいでしょう。

たとえば，国内電話の通話時間は平均2.0分（ほんとうは，もう少し短いのだそうですけれども）の指数分布に従うものとします。ある人が，電話ボックスの前で先客のあくのを待っていますが，もう1.5分も待ったとします。このとき残り時間は平均2.0−1.5＝0.5（分）になるかといいますと，そうではなくて，やはり平均2.0分の指数分布になります。これが記憶喪失型という性質ですが，このような性質をもつものは指数分布だけです。試みに図―2の曲線上の任意の点で接線をひいてみましょう。このとき，接線の$x$軸上への影の長さは，いつも平均の$\frac{1}{\lambda}$になっています。これは指数分布が記憶喪失型であることに起因しているといってよいでしょう。

また，もっと一般的なはなしとしては，次のような考察をされたらよいでしょう。いま，図―2の適当なところで$y$軸に平行な線を立て，密度曲線をちょん切ったとします。そして，残り（右側）の方の面積が1になるように拡大します。すると，拡大した新しい図は，もとの図とピッタリ重ね合わすことのできるものになっています。




図―3　指数分布の性質（2）

図―3は，このことを説明しています。このような性質をもつ密度曲線は指数分布以外にありません。

以上，指数分布のもついろいろな性質を並べあげましたが，それだけでは遊戯にすぎないとお考えでしょう。ところが実は，世の中のさまざまな現象のうち，ドンピシャリ指数分布に従うと考えてよい，というような場合がたいへん多いのです。またドンピシャリとまではいかなくとも，正規分布で近似するよりは指数分布と考えたほうが自然である，というような事例がたくさんあります。むしろ，生データで，それが正規分布に従うなどというのは，よく管理された生産工程での製品の寸法とか，前節で述べた標準的なテストの得点とかのような場面に限定されているようで，社会データから探しだそうとすると，少なからず苦労

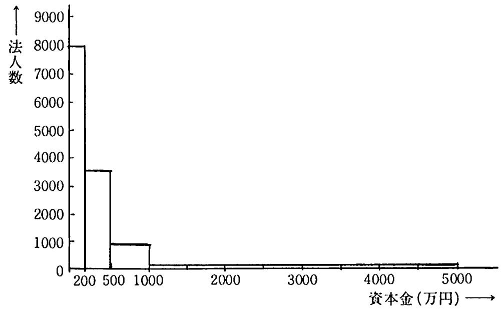
しますが，指数分布的なものは，その気になってお探しになれば，これも指数分布，あれも指数分布という具合に，どこにもコロがっている感じさえしてきます。

お手許にある「統計いばらき」(1978年7月号)をお開きください。8ページに県内法人の規模別分布が載っています。全産業(総数17,014)の資本金規模に応ずる分布は表—1のようになっているそうです。

表—1　県内法人の規模別分布

| 資本金規模 / 産業 | 200万円未満 | 200～500万円 | 500～1,000万円 | 1,000～5,000万円 | 5,000万円以上 | 計 |
|---|---|---|---|---|---|---|
| 全産業 | 8,042 | 5,279 | 2,192 | 1,344 | 157 | 17,014 |

これをヒストグラムに書くときには，資本金規模の階級が不等間隔になっている点に注意して図—4のようにしないといけません。




図—4　県内法人の規模別分布

このことについては9月号でもくわしく述べてあります。
(㊟縦軸の目盛はむしろないほうがよいのですが，ここでは便宜的につけておくだけです)

図—4から，県内法人の規模別分布は指数分布的なものであることがわかります。また同じ8ページに載っている県外法人の従業員数による区分についても，その分布は指数分布とみなしてよいことにお気付きでしょう。

こんどは目先をかえて，デパートの売上に注目してみましょう。毎年6月末の日曜日や12月上旬の日曜日のデパートの繁昌ぶりが新聞紙上などで報じられていますが，ことしの6月は，たとえば日本経済新聞では「中元商戦ことしこそ」という見出しで，個人消費が回復基調にある中での各デパートの売り上げを報道しています。それによると客1人あたりの消費金額は都心で平均4,000円程度のようです。そこで次のようなことを考えてみましょう。都心のデパートでは，6月末の日曜日に「15万人の客がはいって，売上高は6億円に上った」というニュースを聞いた人がいたとします。この人はその日12,000円の買物をしてきたのですが，はたして上位何%程度の客になるでしょうか。

このようなとき，多少でも統計に関心をよせる方ならば，1人あたり4,000円だから，12,000円というのは，かなり上位にあるだろうぐらいのことは，すぐわかります。しかし，ほんとうに明るい人は，12,000円が上位5%以内にはいることを，すぐに見抜きます。それはなぜかといいますと，この種の(売上高の)分布はだいたい指数分布をしているといってよいことを知っているからです。平均が4,000ですから標準偏差も4,000です。12,000という値は平均からちょうど標準偏差の2倍だけ離れているわけですが，それ以上である確率は5%以内であることが，すぐに計算できます。

〔参　考〕

$f(x)=\lambda e^{-\lambda x}$ で表わされる曲線と座標軸の間の面積は1です。(分布関数の面積はいつも1になるようにしてある)そうすると図—2の斜線の部分の面積は，「ある金額$x_1$円までの買い物をしたお客の数が含まれる確率」を表わしています。したがって，ある金額$x$円までの買い物客の数が含まれる割合は

$$F(x)=1-e^{-\lambda x}$$

になります。反対に，ある金額$x$円までの客の数に含まれない割合は

$$R(x)=e^{-\lambda x}$$

となります。

この場合は，$\lambda=\frac{1}{4,000}$　$x=12,000$ を代入して

$$R(12,000)=e^{-\frac{1}{4,000}\cdot 12,000}=e^{-3}$$

指数分布を使うときは，$e^{-x}$の値が必要です。数表になって統計の本に載っています。$x$が3のところを読むと$e^{-x}$が0.049787と出ています。つまり，12,000円を越える買い物をした人の数は約5パーセントの中に含まれることになるわけです。

## ✤まとめ

累積図表の扱い方から始まり，ABC分析，標準偏差と3回にわたって，統計のトピックス的な記事を載せていただきましたが，その中にはかなり高級なセンスも盛り込んでみたつもりです。このシリーズが少しでもお役に立てば幸いです。

## ■統計の窓「お知らせ」

# 第6次漁業センサスの特色

## ── 調査日 11月1日 ──

第6次漁業センサスは，今月1日現在で調査が行われている。しかしながら，その背景は過去の漁業センサスと全く異なっている。すなわち，27年にマッカーサーラインが撤廃されて以来，わが国の漁業は，「沿岸から沖合へ，沖合から遠洋へ」と漁場の外延的拡大を図るなかで発達してきた。しかし，48年に始まった第3次国連海洋法会議において，国際的な合意をみないまま，排他的経済水域を漁業について先取りする形で200カイリ漁業水域を設定する国が相次いでおり，漁業は実質的に200カイリ時代に突入するという沿岸国の海洋分割時代を迎えたことである。

さらに48年の石油ショックを契機として，わが国経済は戦後最大の不況に直面することとなり，漁業においても，燃油および漁業用資材の高騰の下での魚価の低迷によって収益性が低下し，漁業経営は危機感を深め，わが国の漁業経営は，資源浪費型経営から省資源型経営への転換を迫られることとなった。

第6次漁業センサスの特色は以下のとおりである。

**第1は，** 操業水域別の統計作成である。200カイリ漁業水域の設定に伴う入漁，漁業調整，補償等の新たな行政課題に対応するため，漁業経営体および漁船について操業水域を調査し，漁業経営体，漁船および漁船乗組員の操業水域別の実態を明らかにすることである。

**第2は，** 他国の200カイリ漁業水域設定に伴い失業した漁船乗組員の転業または漁業内部の他業種への転換という新たな行政課題に対応するため，自営漁業への還流状況の実態を明らかにする。

**第3は，** 現行漁業制度の見直し等の利用面に応ずるため，従来の漁業種類別統計に漁業制度を結びつけることにより，内容をより一層整備する。

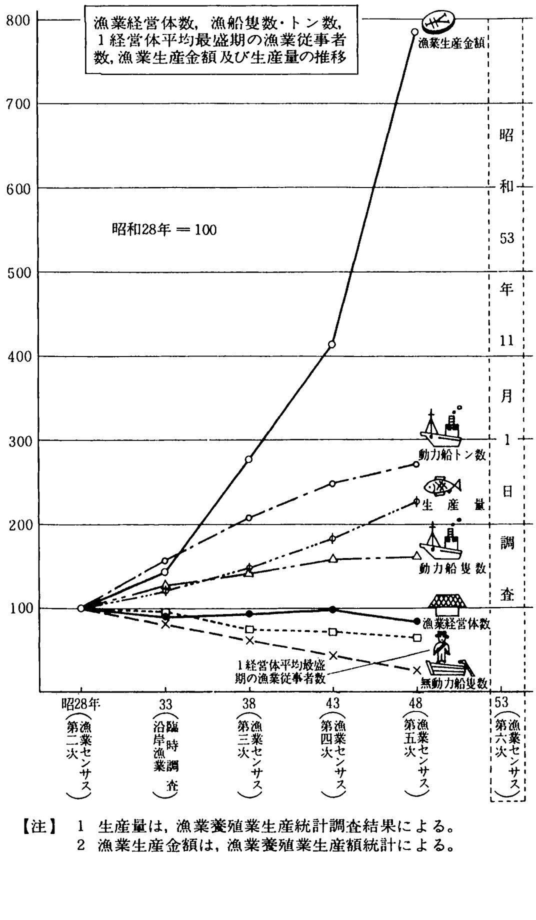
**第4は，** 第5次漁業センサスで把握した個々の経営体と，第6次漁業センサスによる個々の経営体との比較照合を行ない，5年間における新規出現，消滅および継続経営体を識別し，新規と消滅経営体についてはその経営規模・漁業種類を，継続経営体については経営規模・漁業種類・従事日数・労働力等の漁業経営の内容変化を明らかにする。

**第5は，** 沿岸漁場整備開発事業との関連で沿岸漁場環境に関する統計（たとえば幼稚仔の生育場となる藻場の有無，漁場環境悪化の指標として赤潮および油濁発生の有無等）を整備する。

**第6は，** 第6次漁港整備と関連づけた漁村生活環境整備との関連で上下水道，都市ガスの普及状況等，従来より調査していた指標に加え，電話の普及状況等の新しい指標も追加して漁村の生活環境に関する統計を充実する。

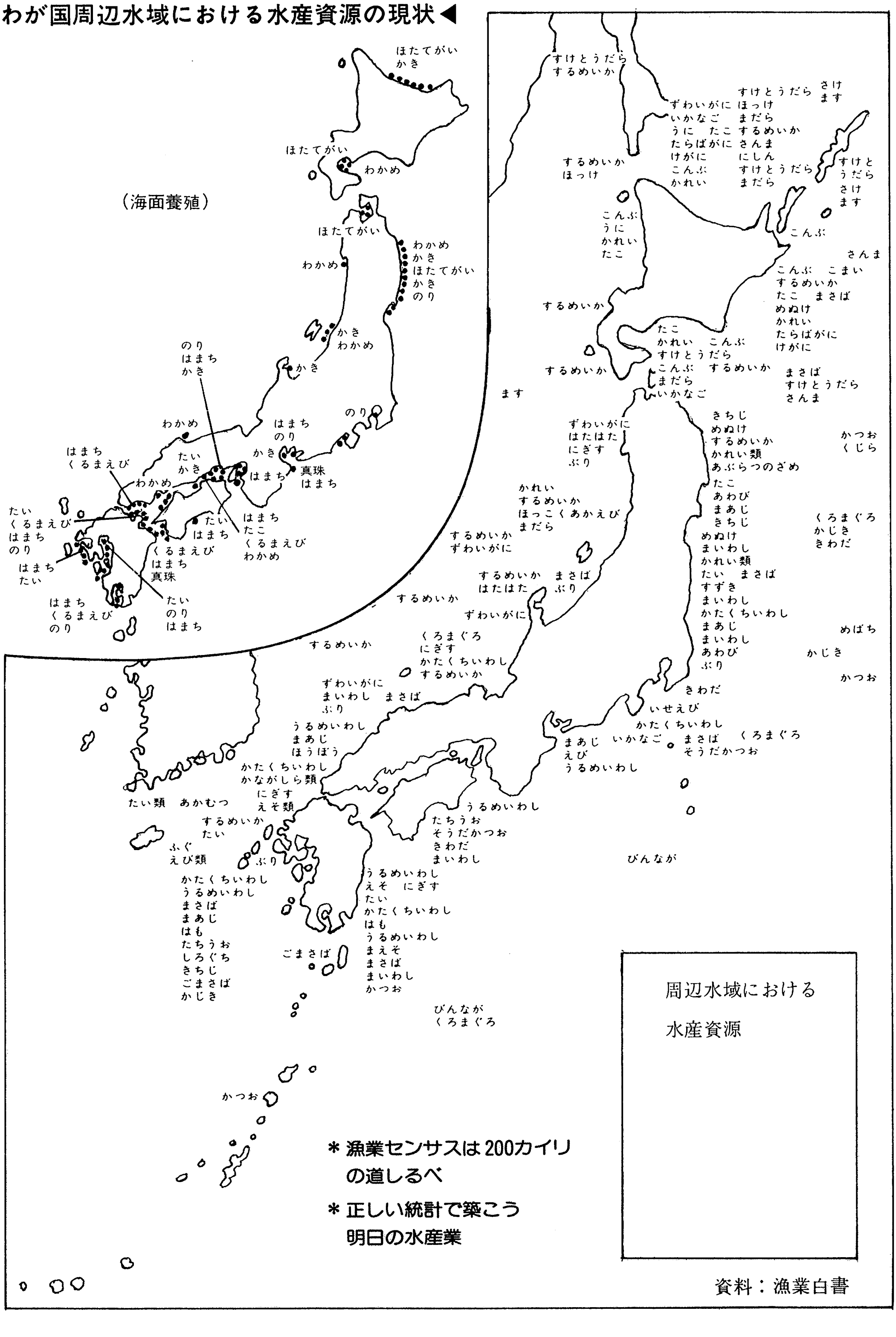
**第7は，** 200カイリ時代を迎えて見直されるようになった内水面養殖業について食用魚に重点を置き，従来はおもな養殖方法別のみの統計であったが，食用魚種ごとに養殖状況を明らかにする統計を作成する。

「農林統計調査 1978.9」より抜粋




【注】 1 生産量は，漁業養殖業生産統計調査結果による。
2 漁業生産金額は，漁業養殖業生産額統計による。

## ▶わが国周辺水域における水産資源の現状◀




資料：漁業白書