

岩波講座 基礎工学 3

確率統計現象Ⅱ

瀧 保 夫 編



岩波書店

1121
101
16-1

岩波講座 基礎工学 3

確率統計現象

II

瀧		保	夫
茅		陽	一
宮	川		洋
関	根	泰	次

3k61/32



岩波講座 基礎工学 3 確率統計現象 II (全 19 巻／第 6 回配本)

1968 年 6 月 26 日 第 1 刷発行 ©

東京都千代田区神田一ツ橋 2-3 株式会社 岩波書店／精興社印刷・松岳社製本

目 次

はじめに

第1章 確率統計現象と確率

1. 1 確率統計現象	1
1. 2 確率の概念	4
1. 3 標本空間と事象	7
1. 4 確率の公理	10
1. 5 結合確率	13
1. 6 条件付確率	15
1. 7 統計的独立	20

第2章 不規則変数と確率分布

2. 1 不規則変数と確率分布関数	23
2. 2 結合確率分布関数	28
2. 3 不規則変数の独立性	32
2. 4 期待値と分散・共分散	33
2. 5 大数の法則	38
2. 6 離散分布	42
2. 7 1次元連続分布	49
2. 8 特性関数	54
2. 9 不規則変数の和	57
2.10 モーメント	60
2.11 多次元連続分布	62
2.12 不規則変数の関数の確率分布	66

第3章 標本抽出

3. 1 母集団と標本抽出	71
3. 2 標本平均と標本分散	77
3. 3 中央極限定理と極限分布	82



3. 4 正規母集団からの標本抽出	91
3. 5 標本の極値	100
3. 6 寿命の分布と信頼性	108

第4章 統計的推定および検定

4. 1 統計的推測と統計的決定理論	115
4. 2 点推定	116
4. 3 推定量	118
4. 4 最尤推定法	121
4. 5 統計的仮説の検定	124
4. 6 尤度比検定法	129
4. 7 統計的仮説の検定の例	136
4. 8 母集団分布によらない検定法	143
4. 9 区間推定	149
4.10 統計的決定理論	152

第5章 最小二乗法

5. 1 回帰関数と最小二乗法	157
5. 2 回帰模型	166
5. 3 標本回帰式	167
5. 4 標本回帰式の分布	172
5. 5 正規回帰	176
5. 6 回帰係数の検定	178
5. 7 分散分析法	184

第6章 不規則過程

6. 1 不規則過程	193
6. 2 定常不規則過程	196
6. 3 相関関数	199
6. 4 周期性不規則過程のフーリエ展開	203
6. 5 パワー・スペクトル	210
6. 6 相互パワー・スペクトル	218
6. 7 濾波と濾波器	219
6. 8 定常不規則過程に対する濾波器の応答	224

6.9 正規性とエルゴード性	228
6.10 マルコフ過程	232
6.11 マルコフ過程における状態遷移確率	238
さらに勉強するために	247

第4章

統計的推定および検定

本章では、標本から母集団の統計的性質について何らかの推測を行なう場合に基礎となる統計的推測の理論と、標本から直接にわれわれの行動を決定する場合の基礎となる統計的決定理論について述べる。まず統計的推測の基本概念について説明し、ついで最尤推定法、尤度比検定法等について述べ、最後に統計的決定理論を紹介する。

4.1 統計的推測と統計的決定理論

母集団からの標本を統計的に整理し、検討する目的は、その整理・検討の結果から、母集団について何らかの推測を行ないたいからである。このようにデータの統計的性質から母集団の性質について推測を行なうことを**統計的推測** (statistical inference) という。統計的推測は**統計的推定** (statistical estimation) と**統計的仮説の検定** (testing of statistical hypothesis) の2種に分類できる。さらにこの前者の統計的推定は**点推定** (point estimation) と**区間推定** (interval estimation) の二つに分けるのが普通である。

このように統計的推測はいろいろの種類に分類して取り扱われているが、最近発展した**統計的決定理論** (statistical decision theory) によれば、これらはすべて**費用関数** (cost function) を最小とするという原則に従って統一して取り扱えることがわかっている。つぎに上記の諸概念を簡単な例を挙げて説明しよう。早朝、勤めに出かけるとき、雨傘を持って行くべきかどうかの問題を考えてみよう。この問題を解くのに費用を考えて解く方法が考えられる。これが統計的決定理論による方法である。この方法では、まず、雨傘を持って行かなかったとき、もし雨に降られれば、どれだけの損失をうけるか、また雨傘を持って行ったとき、もし雨に降られなければ、雨傘を運ぶのにどれだけ費用がかか

るかなどがわかっているものと仮定する。ついで、各種の気象データからその日に雨の降る確率を求め、これらの諸量から、雨傘を持って行ったときに期待される費用と雨傘を持って行かなかったときに期待される費用を計算し、その大小によって雨傘を持って行くか、持っていないかを決定する。

しかしこの統計的決定理論を適用するためには、上に述べたように各種の費用が既知である必要があり、これが未知の場合には、持っていくべきか、行かざるべきかを決定することは不可能である。理想から言えば、雨傘を持って行くべきか、持っていないほうがよいかというように、われわれの行動そのものを直接データから決定するのが望ましいのではあるが、上に述べたようにこれが困難な場合には一步譲って、われわれが行動する場合の参考資料を与えるという立場に立つのが普通である。この立場において必要となるものが、統計的推定あるいは仮説の検定である。たとえば、統計的推定では、今日は何ミリ位の雨量があるかを推定する(点推定)か、あるいは何ミリから何ミリまでの間の雨量があるかを推定する(区間推定)のである。実際の行動はこの推定を参考資料として、われわれの判断にまかせようというのである。また統計的仮説の検定では、今日は雨が降るという仮説と今日は雨が降らないという対立仮説のどちらを正しいとみなしたらよいかを推測する。その結果、どちらを採るにしても傘を持って行くかどうかは、やはりわれわれの判断にまかせられる。

本章では、まず統計的推定と統計的仮説の検定について述べ、最後にこれらの統計的推測の理論と統計的決定理論の関係について述べる。

4.2 点推定

簡単な例として、微小直流電圧値を測定する場合を考えよう。微小電圧を測定する一つの方法に直流増幅器を用いる方法がある。この方法によれば、微小電圧を、まず増幅器によって増幅し、その増幅した出力電圧を通常の電圧計で読みとるか、あるいはペン書き記録計に記録する。この値を増幅器の増幅度で割れば入力電圧が得られる。

電圧が極めて微小なる場合には、図4.1に示すようにこのようにして得られた出力電圧には、求める直流電圧以外に低周波の雑音電圧が重畳するのが普通である。いま $t=0$ から $t=T$ までの出力波形 $f(t)$ が記録されているものとす

る。どのようにすればその中に含まれている直流電圧が推定できるであろうか。この問題の一つの取扱いとして、波形を標本化 (sampling) して取り扱う方法がある。図 4.1 に示すように、0 から T までの区間を $n-1$ 等分し、連続的な波形 $f(t)$ ($0 \leq t \leq T$) をその標本値 $f(0), f(\Delta t), \dots, f((n-1)\Delta t)$ で代表させる。このように標本化し、さらに直流電圧に重畳している雑音電圧は互いに独立に期待値 0 の同一の分散を有する正規分布をなすものとすれば、標本値 $f(0), f(\Delta t), \dots, f((n-1)\Delta t)$ は、期待値が求める直流電圧に等しく、分散は雑音電圧の分散 (雑音電圧の 2 乗平均値) に等しい正規母集団からの n 個の標本値と考えられる[†]。したがって、この問題は、母集団は正規分布をすることはわかっているのであるが、その期待値と分散が未知であるとき、 n 個の標本値よりその期待値を推定せよという問題に帰着する。これは点推定の問題の最も代表的なものであって、分散が未知の正規母集団の期待値の推定という問題である。直流電圧の測定では求めたいものは期待値であるが、雑音電圧の実効値 (雑音電圧の 2 乗平均平方根すなわち雑音電圧の分散の平方根) を求めたい場合もあ

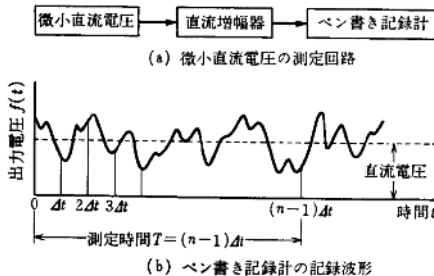


図 4.1 雑音電圧中の直流電圧の測定

[†] 一般に雑音波形の各標本時点の標本値が、その標本時点をどのようにとっても多次元の正規分布をなす場合に、その雑音をガウス性雑音となづける。このようなガウス性雑音の標本値は、一般には期待値が 0 でもなければ同一の分散を有するものでもなく、また相互に独立に分布するものでもない。標本間隔を十分大きくとれば相互に独立に分布するものとみなせるようになるのが普通であるが、実際にどれだけの間隔をとればよいかは雑音の性質によるのであって、簡単には言えない。特にいくら標本間隔をせばめても、標本値が相互に独立に期待値 0 の同一の分散を有する正規分布をなす雑音を白色ガウス性雑音という。これは一つの理想化された雑音であって、実際には存在しないが、直観的には極めて速い、こまかな振動をしている雑音ともいえる。このような雑音に対しては、本文中の仮定が標本空間のいかに問わず成立する (詳しくは第 6 章参照)。

る。このときは母集団の分散を推定する必要がある。

このように点推定は、母集団分布に関する何らかのパラメーターが未知の場合に、標本抽出により得られる母集団の標本値を用いて、このパラメーターの値を推定することを目的とする。このとき母集団分布に関するパラメーター θ を母数といい、標本値から推定された値を推定値 (estimate) といい、 $\hat{\theta}$ であらわす。一般に推定値 $\hat{\theta}$ は標本値 $x_1, x_2, \dots, x_n$ から推定するのであるから、

$$\hat{\theta} = g(x_1, x_2, \dots, x_n) \quad (4.1)$$

とあらわされなくてはならない。しかるに $x_1, x_2, \dots, x_n$ は母集団と同一分布を有し、かつ相互に独立な不規則変数 $X_1, X_2, \dots, X_n$ の実現値であるから、 $\hat{\theta}$ は同じく対応する統計量

$$\hat{\Theta} = g(X_1, X_2, \dots, X_n) \quad (4.2)$$

の実現値と考えられる。このように推定値 $\hat{\theta}$ を実現値としてもつ統計量 $\hat{\Theta}$ を、母数 θ の推定量 (estimator) という。すなわち推定量は一つの不規則変数である。

4.3 推定量

前節の直流電圧の推定の問題をもう一度考えてみよう。いま図4.1のように T 秒間の波形から等間隔に n 個の標本値が得られたとし、これらの標本値を順番に $x_1, x_2, \dots, x_n$ とする。すでに述べたように、統計的推定では推定値 $\hat{\theta}$ は、(4.3)のようにこれらの標本値 $x_1, x_2, \dots, x_n$ の関数としてあらわされる。

$$\hat{\theta} = g(x_1, x_2, \dots, x_n) \quad (4.3)$$

ここで、 $x_1, x_2, \dots, x_n$ はそれぞれ不規則変数 $X_1, X_2, \dots, X_n$ の実現値であるから、 $\hat{\theta}$ も、つぎのような推定量 $\hat{\Theta}$ の一つの実現値にほかならない。

$$\hat{\Theta} = g(X_1, X_2, \dots, X_n) \quad (4.4)$$

さて、さきに述べた直流電圧の測定の問題では、標本値 $x_1, x_2, \dots, x_n$ は同一の正規母集団からの、互いに独立な標本値であり、推定すべき母数は、その母集団平均である。3.2節で述べたようにデータには母集団の性質が反映されているはずである。ところでいま推定しようとしている母集団平均は、分布の中心的な傾向を示す量であり、当然これはデータの中心的な傾向にあらわれているものと考えられる。データの分布の中心的な傾向を示す量としては、標本値

の単純平均すなわち標本平均，中央値(median)あるいは最頻値(mode)などいろいろの量が考えられる。これらの値はいずれも標本値 $x_1, x_2, \dots, x_n$ の関数であり，したがって(4.3)に示すように統計的推定における推定値として資格を有する。

上に述べたように，推定量としては，いろいろな量が考えられる。どれが最も望ましい推定量であるかがつぎに問題となる。その一つの考え方として，計算しやすさというものがある。たとえば $\hat{\theta}_1$ という推定量と $\hat{\theta}_2$ という推定量が存在し，両者とも同じような性質を持っているが， $\hat{\theta}_1$ は $\hat{\theta}_2$ に比べて極めて計算が容易であるというような場合である。両者ともほぼ同一の値を与えるなら，計算しやすい推定量が望ましいことはいうまでもない。しかし，普通は $\hat{\theta}_1$ は $\hat{\theta}_2$ より計算しやすいが，推定量としては $\hat{\theta}_1$ のほうが $\hat{\theta}_2$ より平均として真の母数から離れているというような場合が普通である。データの数に極めて多い場合には，多少推定量としては劣っていても計算しやすい $\hat{\theta}_1$ を採用することもまれではない。このような方法は多くの場合，母集団の形によらない，すなわちノンパラメトリックな方法(non-parametric method)になるのが普通である。これについては後節で詳しく述べることにし，ここでは，計算の容易さは問わないこととして，一般的に推定量としては，どのような性質を有するものが望ましいかを考えてみよう。

推定量 $\hat{\theta}$ は一つの不規則変数である。その密度関数は $X_1, X_2, \dots, X_n$ のしたがう分布に依存する。たとえば，すでに述べた直流電圧の測定では，推定量 $\hat{\theta}$ は測定値 $X_1, X_2, \dots, X_n$ の関数であり， $X_1, X_2, \dots, X_n$ が正規分布にしたがうものと仮定すれば， $\hat{\theta}$ のしたがう確率分布は $X_1, X_2, \dots, X_n$ のしたがう正規分布の期待値 μ と分散 σ^2 に依存する。直流電圧の測定では正規分布の期待値 μ が未知であり，それを推定する。いまこの推定量として， $\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3$ と三つ異なるものがあり，これらは真の母数 θ のまわりに図4.2に示すように分布しているものとする。

この図の $\hat{\theta}_1$ あるいは $\hat{\theta}_2$ のように，その期待値が真の母数 θ と一致するものを不偏推定量(unbiased estimator)という。また

$$E[\hat{\theta}] = \theta \quad (4.5)$$

なる性質を不偏性(unbiasedness)とよぶ。これは推定量として望ましい一つの

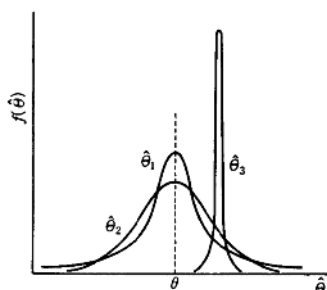


図 4.2 各種の推定量の分布

性質である。これに対し $\hat{\theta}_3$ の期待値は θ に等しくない。このような推定量を **偏推定量 (biased estimator)** という。しかし偏推定量はつねに不偏推定量より劣る推定値を与えるというわけではない。たとえば図 4.2 では $\hat{\theta}_1$ は不偏推定量ではあるが、個々の推定値を考えてみると、 $\hat{\theta}_1$ は場合によっては $\hat{\theta}_3$ とは異なり真の推定値から極めてはなれた値を与えることもある。ただ、不偏推定量は、同様の測定を何度も繰り返し、推定値が多数得られた場合には、これらを平均することにより、大数の法則で、いくらでも真値に近い値を得ることができるという特徴を有する。これが推定量に、不偏性という性質が望まれる理由である。

これと関連して、標本の大きさ n が十分大きくなった場合に、推定量 $\hat{\theta}$ が真値 θ より、任意に選ばれた ε 以上離れる確率が 0 に近づく場合、この推定量を **一致推定量 (consistent estimator)** とよび、この性質

$$\lim_{n \rightarrow \infty} \mathbf{P}(|\hat{\theta} - \theta| < \varepsilon) = 1 \quad (4.6)$$

を **一貫性 (consistency)** という。大きさ n の標本 $X_1, X_2, \dots, X_n$ に対して、不偏推定量 $\hat{\theta}(X_1, X_2, \dots, X_n)$ があるものとする。いま大きさ nm の標本があった場合、これを $X_1, X_2, \dots, X_{nm}$ としたとき、

$$\begin{aligned} \hat{\theta}_0(X_1, X_2, \dots, X_{nm}) \\ = \frac{1}{m} \sum_{i=0}^{m-1} \hat{\theta}(X_{in+1}, X_{in+2}, \dots, X_{(i+1)n}) \end{aligned} \quad (4.7)$$

なる推定量 $\hat{\theta}_0$ を考えると、大数の法則により、 m が十分大きくなったとき、

$\hat{\theta}_0$ が θ より任意に選ばれた ε 以上離れる確率は 0 に近づくことが示される。すなわち、このような意味で不偏推定量は、一致推定量の特別なものであることがわかる。したがって不偏推定量が得られない場合、それに代わる推定量として一致推定量を用いるのが普通である(たとえば、次節で述べるように最尤推定量は必ずしも不偏ではないが、極めて一般的な条件の下で一致推定量である)。

つぎに図 4.2 の $\hat{\theta}_1$ と $\hat{\theta}_2$ のような二つの不偏推定量同士の優劣について考えよう。この図で $\hat{\theta}_1$ は大部分は θ のまわりに分布しているが、ときどき極めて θ から離れた値を与える。これに対して $\hat{\theta}_2$ は $\hat{\theta}_1$ にくらべて大部分のものは $\hat{\theta}_1$ よりもばらつきは大きい、極めて大きく離れた値を与えることはない。このような二つの不偏推定量の優劣を比較するには、これらの推定量の分散が用いられることが多い。特に最も分散の小さな不偏推定量を有効推定量 (efficient estimator) とよぶ。また有効推定量の分散と、他の不偏推定量の分散の比を、この後者の推定量の効率 (efficiency) とよんでいる。

このように分散を用いて二つの不偏推定量の比較をすることができる。一般の推定量、すなわち偏移推定量まで含めて、その優劣を比較したい時には、誤差の 2 乗平均を用いるとよい。これは

$$\begin{aligned} E[(\hat{\theta} - \theta)^2] &= V[\hat{\theta}] + \{E[\hat{\theta}] - \theta\}^2 \\ &= V[\hat{\theta}] + (\text{偏移量})^2 \end{aligned} \quad (4.8)$$

を用いる方法である。この規準を二乗平均誤差規準 (mean square error criterion) とよぶ。この値の最も小さな推定量を最小二乗平均誤差推定量 (least mean square error estimator) とよぶ。特に不偏な最小二乗平均誤差推定量が有効推定量にほかならない。

4.4 最尤推定法

前節で、推定量に望まれる性質について述べたが、このような望ましい性質を有する推定量を実際に $X_1, X_2, \dots, X_n$ の関数として求めることは、必ずしも容易ではない。つぎに、逆に必ずしも望ましい性質を有するとは限らないけれども、比較的簡単に導くことのできる最尤推定量について、それがどのような性質を有しているかを考えてみよう。

いま、母集団からの大きさ n の標本を $x_1, x_2, \dots, x_n$ とし、母集団のしたがう確率密度関数を $f(x; \theta)$ とすれば、標本 $x_1, x_2, \dots, x_n$ が独立なものであれば、その結合確率密度関数は次式で与えられる。

$$f(x_1, x_2, \dots, x_n; \theta) = f(x_1; \theta)f(x_2; \theta) \cdots f(x_n; \theta) \quad (4.9)$$

ここで θ は母数を示す。

ここで標本 $x_1, x_2, \dots, x_n$ は既知の値であり、母数は未知であるとすれば、(4.9) は θ の関数と考えられる。このように大きさ n の標本の結合確率密度関数を θ の関数と考えたとき、この関数を尤度関数 (likelihood function) とよび L で表わす。すなわち

$$L = f(x_1; \theta)f(x_2; \theta) \cdots f(x_n; \theta) \quad (4.10)$$

いま標本値 $x_1, x_2, \dots, x_n$ が与えられた場合、(4.10) の L を最大とするような θ の値 $\hat{\theta}$ を最尤推定値 (maximum likelihood estimate) とよび、このような $\hat{\theta}$ を与える統計量 $\hat{\theta}$ を最尤推定量 (maximum likelihood estimator) という。またこのような推定法を最尤推定法 (maximum likelihood estimation) という。

この最尤推定量は必ずしも不偏推定量を与えるとは限らない。しかし、極めて一般的な条件の下で、最尤推定量は一致推定量であり、しかも標本の大きさ n が大きくなると、漸近的に正規分布をなすと共に、また漸近的に有効推定量となることが証明されている[†]。また n が有限の場合は、必ずしも不偏ではないが、これは簡単に補正できるので、実用上は有効推定量と同様に取り扱うことができ、しかも、 $X_1, X_2, \dots, X_n$ に関する関数形が簡単に得られるのが特徴となっている。

未知の母数がただ一つでない場合には、最尤推定法ではこれらのすべての母数について L を最大とするようにする。したがって母数を何個か含むような場合には、たとえ我々が求めたい母数がそのうちの一つに限られていても、場合によって他の未知母数の最尤推定値を同時に求める必要が起こることがある。つぎに例を挙げ、これを説明しよう。

いま $x_1, x_2, \dots, x_n$ を、期待値 μ が未知の正規母集団からの標本値であると仮定する。このとき尤度関数 L は

[†] これらに関する詳細は Cramer, H.: *Mathematical Methods of Statistics*, Princeton Univ. Press(1951) §33.2-33.3 を参照のこと。

$$L = \left(\frac{1}{2\pi\sigma^2} \right)^{n/2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right\} \quad (4.11)$$

となる。つぎに L を最大ならしめる μ の値を計算する。計算を簡単にするために、 L そのものでなく、 $\log L$ を最大ならしめることを考えよう。一般に L の対数を対数尤度関数とよんでいる。

$$\log L = -\frac{n}{2} \log (2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \quad (4.12)$$

(4.12) を μ について偏微分すれば、

$$\frac{\partial \log L}{\partial \mu} = -\frac{1}{2\sigma^2} \sum_{i=1}^n 2(x_i - \mu)(-1) \quad (4.13)$$

(4.13) を 0 とすれば、

$$\sum_{i=1}^n (x_i - \hat{\mu}) = \sum_{i=1}^n x_i - n\hat{\mu} = 0 \quad (4.14)$$

すなわち

$$\hat{\mu} = \frac{\sum_{i=1}^n x_i}{n} \quad (4.15)$$

が得られる。すなわち正規母集団の母集団平均の最尤推定値は標本平均にほかならないことが導かれた。

つぎに母集団分散の最尤推定値を求めてみよう。同様に対数尤度関数を σ^2 について偏微分すれば、

$$\frac{\partial \log L}{\partial \sigma^2} = -\frac{n}{2} \frac{1}{\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (x_i - \mu)^2 \quad (4.16)$$

となる。したがって最尤推定値はこれを 0 とおいて、

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (x_i - \mu)^2}{n} \quad (4.17)$$

となる。 μ が既知の場合は (4.17) より $\hat{\sigma}^2$ の推定値が得られる。 μ が未知の場合は、(4.15) により μ の最尤推定値 $\hat{\mu}$ を求め、これを (4.17) に代入する。すなわち標本平均を $\bar{x}$ とすれば

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} \quad (4.18)$$

が、平均値が未知の正規母集団の母集団分散の最尤推定値である。(4.18)は σ^2 の最尤推定値であるが、 $\log L$ を σ について偏微分して、それを0とおけば、 σ の最尤推定値が得られる。容易に計算によって確かめられるように

$$\hat{\sigma}^2 = \sigma^2 \quad (4.19)$$

が証明できる。

この性質は $\hat{\sigma}^2$ と σ とに関する特別な関係ではなく、最尤推定値の一般的な性質である。すなわち一般に $f(\theta)$ を θ の1価関数とし、 $\hat{\theta}$ を θ の最尤推定値とすれば、 $f(\theta)$ の最尤推定値 $f(\hat{\theta})$ は $f(\hat{\theta})$ に等しいことが証明できる。

$$f(\hat{\theta}) = f(\hat{\theta}) \quad (4.20)$$

これを最尤推定値の不変性(invariance)という。

さて、ここで(4.18)と(3.6)を比較すれば、母集団分散の最尤推定値 $\hat{\sigma}^2$ は実は標本分散 s^2 にほかならないことがわかる。したがって $\hat{\sigma}^2$ を実現値とするような不規則変数 $\hat{\Sigma}^2$ の期待値は、(3.18)に示すように

$$E[\hat{\Sigma}^2] = \frac{n-1}{n} \sigma^2 \quad (4.21)$$

である。このように平均値が未知の場合は正規母集団の母集団分散の最尤推定量 $\hat{\Sigma}^2$ は不偏推定量ではない。これに対し、3.2節に示したように不偏分散

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \quad (4.22)$$

は、母集団分散の不偏推定量である。

4.5 統計的仮説の検定

いま遠距離レーダーの受信信号を増幅したところ図4.3のような波形が得られたものとする[†]。レーダーの場合には、この受信波形の中に反射波が存在するか否かが問題である。図4.3で波形が不規則な変動を示しているのは、雑音

[†] 通常のパルスを用いたレーダーでは受信信号はパルス状である。しかし、連続波を用いるレーダーもあり、この種のレーダーでは、受信信号は図4.3に示すように連続波となる。

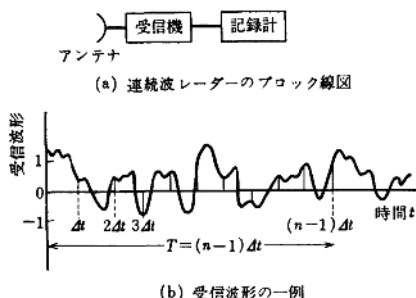


図 4.3 連続波レーダーの受信波形

電圧の影響を示すものであり、雑音がなく、反射波のみの場合には受信波形は直流の一定値(正の一定電圧のこともあるし、負の一定電圧のこともある)を示す。さて $t=0$ から $t=T$ までの受信波形を観測し、この結果から反射波が存在するかどうかをどのようにすれば検定できるであろうか。

ここではこの問題を 4.2 節の直流電圧の測定と同様に、波形を標本化して取り扱うことにする。すなわち図 4.3 に示すように $0 \sim T$ までの区間を $n-1$ 等分し、連続的な波形 $f(t)$ ($0 \leq t \leq T$) をその標本値 $f(0), f(dt), \dots, f((n-1)dt)$ で代表させる。さらにこの波形に重畳している雑音電圧の標本値は互いに独立に期待値 0 の、同一の分散を有する正規分布をなすものと仮定しよう。しかれば受信波形の標本値は、もし受信波形中に反射波が含まれていなければ、直流分がないのであるから、その期待値は 0、分散は雑音電圧の分散(すなわち雑音電圧の 2 乗平均値)に等しい正規母集団からの n 個の標本値と考えられる。したがって上記の受信波形中に反射波が含まれているかどうかを検定する問題は、 n 個の標本値 $f(0), f(dt), \dots, f((n-1)dt)$ が母集団平均 0 の正規母集団からの標本値とみなせるか(これは受信波形が雑音のみであることに対応する)、あるいはある 0 以外の母集団平均を有する正規母集団からの標本値とみなさなければならないか(これは受信波形に雑音以外に反射波が含まれていることに対応する)のどちらを採用すべきであるかを検定する問題になる。これは統計的仮説の検定の代表的なものであって、正規母集団の母集団平均に関する検定といわれているものである。