確率は確信の度合い

原　宣一

はじめに

確率について皆さんにお話したいことがあります。それは大げさに言えば、私が３０年に及ぶ仕事に関連して気がついた大発見なのです。

最初に気がついたのはまだ２０代の頃でした。それは、現行の確率理論はおかしいのではないかという疑問でした。確率を使って正しく表現する方法に納得できないものがあったのです。

最終的に得た結論は、現在工学関係で応用されている確率理論は基礎が間違っているということなのです。私の恩師の一人近藤次郎先生が「応用確率理論」を書かれています。この本も含めて、確率の基礎、つまり確率の定義が間違った法を採用しています。私は近藤次郎先生にも恐る恐る申し上げてあります。

私が面白いと思うものは世間がひっくり返るような話です。世間の常識であったことがひっくり返されるような話です。この意味で歴史上最も面白い話は天動説から地動説への転換でしょう。「コペルニクス的転換」という用語まで出来ています。アインシュタインの光速度一定の原理もそうでしょう。

このような面白いことはそれほど頻繁にあるわけではありません。でも、どの分野にもときどきあるようです。

いずれも最初は異説として退けられています。もちろん、ほとんどの常識に反する異説は単なる誤解や錯覚であったりすることが多いのです。異説は異説のまま埋もれていきます。

例えば、異説の例として、「相間本」というものがあります。「相対性理論は間違っている」という趣旨のことが書かれた本のことですが、この異説の論拠がおかしいことは私でも指摘できます。

しかし、説得力のある異説はひょっとすると本当ではないかという期待があります。残念ながら、専門分野の違う話で

すとどちらにも軍配を上げられないだけです。

このような例の一つが環境問題です。「環境問題はなぜウソがまかり通るのか2」にいくつかの事項が間違っていると指摘されています。著者の武田邦彦教授の言い分には十分説得力があるように私には思われます。

私がこれからお話する確率の話も、工学分野ではまだ異説に属するでしょう。残念ながら少数派なのです。私が周りの人にいくら説明しても判ってくれません。しかし、反論をもらう訳でもないのです。私は何故そのように無視されてしまうのか考えてみました。

判ったことは殆どの人が、判断の根拠に現在の権威ある説にあっているかどうかで判断しているということです。宇宙ですと、すぐＮＡＳＡが引き合いに出されます。

私は、現行の応用理論がおかしい、ＮＡＳＡも気がついていないと言っているのですから、あえて現在の常識・権威に逆らった話をしているのです。自分の頭で考えて貰わないと私が言っていることが正しいことの軍配は下せないでしょう。

私の世代ではどうしても既存の知識が却って邪魔をしているとしか思えません。そこで柔軟な頭の持ち主である筈のまだ若い学生さんたちに聞いてほしいと思った次第です。

「相間本」の内容のような、間違った話を聞かされるのではないかと思った方、安心してください。

私も正直なところ7，8年前までは私の方が間違っているのではないかとの不安が少しはあったのですが、物理学者の書いた本を見つけてからは、確信を持って話しています。

その物理学者が間違っていたらどうするのかと思われる方もおられるでしょう。当然です。

それでは、もっと確実なことを言いましょう。私の言う正しい確率は、現代物理学で採用されている確率であり、最近目覚しく発展中の情報工学での確率理論とも一致するのです。これほど確かなことはないでしょう。

さて、４０年近く昔に戻って、確率理論の応用がどうもおかしいと思うようになった頃です。私は大きな書店の確率論の棚をあちこち覗いてみました。そしてベイズ流統計学というものもあることを知りました。

「ベイズ流」という断りがあること自体が本流でないことを示して居るみたいです。しかし、私はこちらの方が正しいと思い、ベイズ流統計学による一つの論文を書いて、日本航空宇宙学会に投稿しました。ワープロのない時代ですから手書きの論文です。投稿結果は「理解できない」との一言添えられて返却されてしまいました。誰が査読をしたのか容易に見当はつきました。日本では同じテーマで論文を出している人はいなかったからです。私の論文の内容がその人の方法を否定していたから、「理解できない」のも当然です。

会社の上司の薦めもあって構造強度に関するシンポジウムで発表させて貰いました。

しかし、大方の人の興味を引くことは出来ませんでした。内容は、航空機の安全寿命の決め方に用いる安全係数、これは「ばらつき係数」と呼ばれていますが、この安全係数の合理的な決め方に関するものでした。ばらつき係数の話は、私が確率に疑問を持ったきっかけですので、いつか詳しくお話したいものですが、これだけでも長くなってしまいますので本日は省略します。

その後もずっと確率のことは出てくるたびに気にはしていたのですが、ＮＡＳＡも間違えているという説明はまず納得して貰えなかったのです。

何がおかしかったかと言うと、突き詰めると確率の定義が良くなかったということとそれにつきるのです。

それでも、なかなか納得して貰えないのです。学者同士が７０年間も論争を続けた過去の経緯もあります。

1．確率の例

皆さん丁半博打を知っていますか。江戸時代のやくざもの映画でよく出てきますが、やくざがお客を集めて寺で博打をするときの話です。サイコロ二つを碁笥に入れて振り、二つのサイコロの出目の合計が偶数か奇数かを当てるだけの単純なゲームです。

偶数を丁と呼び、奇数を半と言ったのです。

大事なことは「碁笥」に入れて振るということです。ガラスのコップに入れて振ったのでは駄目です。中が見えるからです。今日ここに持参しましたものはラスベガスあたりで使っている壷で革張りです。サイコロ5個と一緒にアンティーク・ショップで見つけて買っておいたものです。

そして、大事なことのもう一つは「丁」か「半」を予想して振るのではなく、振った後に結果を当てるのです。これはいかさまを少しでも防ぐためにそうするのですが、確率は時間の概念に無関係だということが重要です。

「事象の発生確率が」などという風に言いますが、確率はその事象が発生したか、これから発生するであろうことか。つまり時制は関係ないのです。

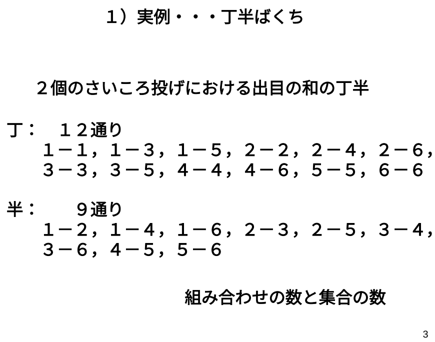
出目の組み合わせは丁が1と1、1と3など全部で１２通りあります。半になるのは1と2、1と4など全部で9通りしかありません。このため、昔の人は丁の方が半より出やすいと考えた人も多かったようです。

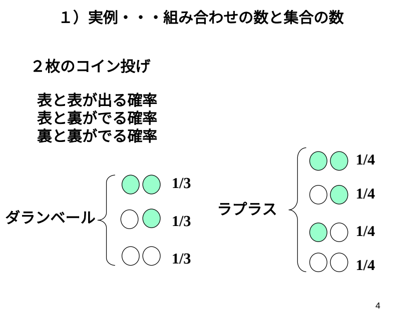
皆さんは丁も半も同じ程度に出やすいと考えるでしょう。でも何故同じと考えるのでしょうか。何故丁も半も確率は0.5なのでしょうか。考えてみたことはありますか。

2個のサイコロ振りよりも簡単なゲームは2個のコインを

振ることです。英語ではコインを振って表がでるか裏がでるかといわずに、頭か尻尾かというのです。西洋人の感覚を疑いますね。

２個とも表が出る確率はどうでしょうか。数学者のダランベールは２個とも表、２個とも裏、１個が表で１個が裏の３通りの出方があるから、２個とも表がでる確率は１／３であるとしました。





ダランベールの弟子であったラプラスは１個のコインが表と裏がでる確率が独立に１／２の場合は、２個のコインを区別すべきで、出方は４通りある。２個とも表が出る確率は１／４とすべきだとしました。

これは実際に何度も実験して統計を取ってみますと、２個とも表が出る割合は１／４に近くなることが多いのです。皆さんは当然のこととお考えでしょうか。

次の例はもう少し難しくなります。

西洋の賭博師でシュバリエ・ド・メレという人がいました。この人は有名なカジノの賭博師で、数学に名を残しているのです。自分がやっているサイコロ賭博の問題を数学者に聞いたからです。賭博師でありながら数学に名を残している人はこの人ぐらいでしょう。大事なことは、何が問題かに気づくことです。重要な問いかけであれば、答えは他の人に任せても名を残せるということです。

最初の問題は、「さいころ1個を４回振ると1度は6が出る」ということに賭けたほうが良いかどうかです。つまり確率が０．５より大きいか否かと言う問題です。

２番目はもう少し複雑で、「サイコロ2個を同時に振るという実験を２４回繰り返すと、一度は2個とも6が出る」は確率０．５以上か、否かという問題です。

これら二つの例は純然たる確率計算です。サイコロ一つの目がでる確率がすべて１／６であるならば、独立事象の合成確率は演繹的に解けます。どのように解くかは少し頭を使う必要があります。前者の問題では一度も6が出ない確率を計算して１からその値を引いたものが答えです。１マイナス（５

／6）の4乗は0．52ですから賭けた方が良いのです。後者の問題は同じように計算すると、0．49ですから賭けないほうが良いことになります。

次に紹介する問題は、正しい答えが多くの人の直感に反することが多いので重要です。「確率とデタラメの世界」という本からの引用です。

「目撃情報の信憑性」と標題をつけてあります。

ある町で起こったタクシーのひき逃げ事件です。その町ではタクシー会社が二つあって、A社はその町の85％のタクシーを占めています。B社のタクシーは15％です。そしてA社のタクシーの色はグリーンで、B社のタクシーはブルーに塗られています。ひき逃げ事件が起きたのは夕方でもう日が暮れかけていました。

幸い目撃者がいて、その目撃者が「ひき逃げした車の色はブルーであった」と証言したのです。

このような夕方では目撃の信憑性が絶対的なものではありません。調べてみますと、目撃情報の正しさは80％であることがわかりました。

さて、ひき逃げした車がB社の車であった確率とA社の車であった確率はどちらが大きいかというのが問題です。

目撃の信憑性は絶対的ではないけれども、80％は正しいのです。B社の車であった確率が、A社の車であった確率より高い、と多くの方が考えます。皆さんはどのように思いますか。

80％信憑性がある目撃者の証言によりA社の車である確率が高いだろうと考えた人は、A社の方が使っている車の数が多いという事前情報を無視しています。利用可能な情報を無視すると間違った答えに結びつくことが多いのです。

正しくは次のように考えなければなりません。

この町のタクシーは全部で100台であったとします。するとA社の車は85台でB社の車は15台です。目撃者の信憑性は80％なので、ひき逃げ車がA社であった可能性は見間違えた可能性が20％ありますので85台×20％は17台。これに対して、B社の車であった可能性は15台×80％で12台です。従って、ひき逃げした車がA社であった確率は17／29でA社の方が高いのです。

このように考える方が合理的なのです。

実例の最後にもう一つ面白いゲームを紹介します。モンティホール問題という名前がついています。

モンティホールというのは人の名前で、テレビの司会者です。日本で言えば「みのもんた」でしょうか。

アメリカの視聴者参加番組は昔から賞金額が大きかったのです。参加した視聴者が賞金や賞品を貰って大喜びをする画面を写して視聴率を上げる戦略なのでしょう。日本の場合は

賞金額の上限が低く抑えられていましたのでそういうことは出来ませんでした。

　このゲームはスタジオに三つのドアがあるのです。そしてどれか一つのドアの後ろにはキャデラックが置いてあり、そのドアを開けたら、キャデラックを差し上げますというものです。

　他の二つのドアには山羊が繋がれているのです。何故山羊かはキリスト教の影響でしょう。キリスト教では羊は良くて山羊は可愛そうな扱いを受けています。何故か知りませんが。

　ただ三つのドアから一つを選ばせるだけでは何も面白くありません。

　スタジオに来たゲストはまず一つのドアを選ばせられて、その前に立ちます。するとモンティホールは残った二つのドアの内、少なくともどちらか一つは山羊ですから、山羊の方のドアを開けるのです。

　そして、「あなたはこのドアを選ばなくて良かった。さて、あなたは残った一つに変えても良いですよ。」と言うのです。

　このとき、あなたならどうしますか。大抵の人は車が当たる確率は３３％から５０％に増えた。もし変更して外れた場合いには変えたことが残念で堪らないだろうからとの理由で変えない人が多いのです。

　しかし、正しい答えは変えなければ車が当たる確率は３３％のままなのです。ドアを変えれば当たる確率は６７％になります。何故か判りますか。

１）実例・・・目撃情報の信憑性

タクシーのひき逃げ事件

タクシー会社A：車の台数　８５％　　グリーンの車
タクシー会社Ｂ：車の台数　１５％　　ブルーの車

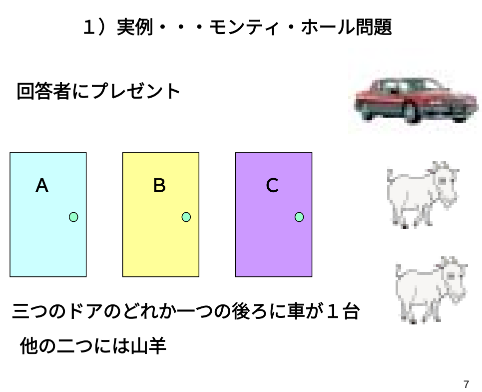
ひき逃げ目撃者はブルーの車と証言

目撃の正しさ：　８０％

ひき逃げした車がタクシー会社Ｂである確率はどれだけか

「確率とデタラメの世界」より

6



確率が５０％に増えたと考えるのは最初にドアが３つあったという事前情報を無視したことになるのです。

　自分が選んだドアとその他の２グループと考えれば自分の選んだドアは３３％の確率だったのです。残りの６７％のうち、モンティホールが山羊の方を開けてしまったので、残った二つのうちの一つのドアが当る確率は６７％であるというわけです。

2. 確率の定義

確率の定義にいろいろあると聞いて「知らなかった」という人が大勢います。皆さんの中にもいらっしゃるでしょう。

世の中の本を見渡すと、現在大きく分類して４種類あります。一番古いものがラプラスによって定義されたもので、起こり得る全ての状態とその中で好ましい状態の比率とするものです。

次にフォン・ミーゼスは相対頻度の極限値を確率と定義しました。

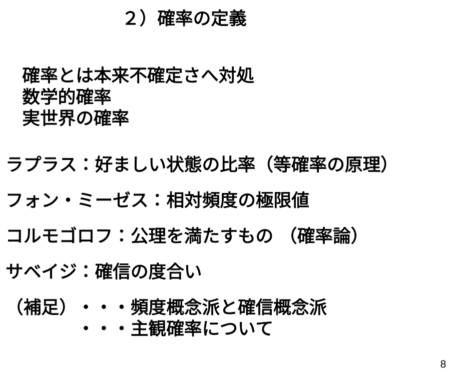
ラプラス流の確率定義とフォン・ミーゼスの頻度確率はそれぞれの学派が優劣を主張し激しい議論がなされました。

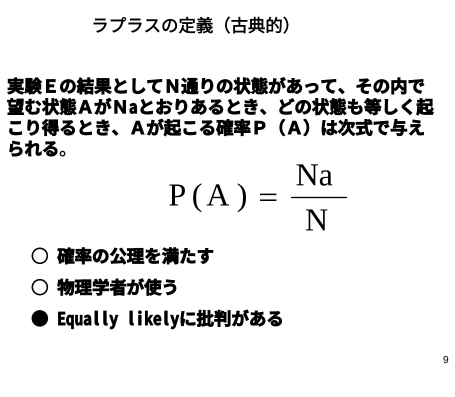
数学者コルモゴロフは公理に基づく確率理論を作りました。数学者は公理を満たすものを確率としたわけです。その確率が実世界で何を意味しようと彼らは興味がないのです。

頻度概念の確率理論はラプラス流の確率を古典確率に追いやる程に優勢であったのですが、サベイジ等により確信の度合いを確率とする定義が出され論争は続きました。

確信の度合いを確率とするものはベイズ流統計学でも採用されている確率です。１９６０年頃、米国を中心にベイジアン革命と言われるほどの動きがあって、多くの分野で頻度概念確率は追いやられました。しかし、分野によっては今も使われています。情報工学を除く工学関係も頻度概念確率です。

それではこれらがどう違うのか、もう少し詳細に説明いたします。





ラプラスの定義

最初がラプラスの定義です。これは次のように定義したものです。

「実験Ｅの結果としてＮ通りの状態があって、その内で望む状態Ａが Na 通りあるとき、どの状態も等しく起こり得るとき、Ａが起こる確率 P(A)は Na/N である。」

この確率の定義でもコルモゴロフの公理を満たします。そ

してこの定義は物理学で使われている確率でもあります。

この確率では問題が多いと考えられた理由は、「等しく起こり得る（equally likely）という表現にありました。等しく起こり得るかどうか判らないではないかという不満です。

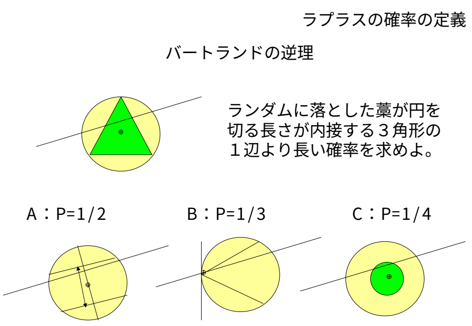
ケインズによる「等確率の原理」は、ラプラスの定義による確率をより基礎的なものに言い換えただけですから同じと見て良いでしょう。等確率の原理の方が明瞭な定義だと思いますが時代が残す名前とは必ずしも実際とは合っていません。ラプラスの定義自体、ラプラス以前に誰かが使っていたものと伝えられています。

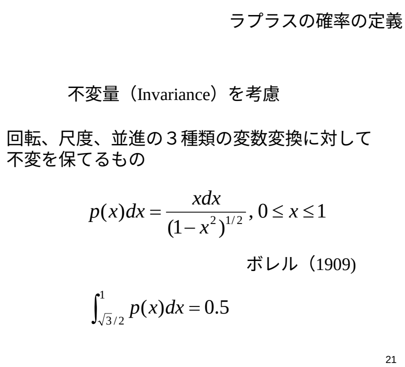
ラプラスの確率が排斥されたとき、もう一つの理由として色々なパラドックス(逆理)が生じるからだとされたのです。しかし、その多くが安易に無限の概念を使ってしまっていたからです。パラドックスで一番難解であったものは、バートランドの逆理と名付けられた次の問題です。

「平面上に円とその円に内接する3角形が描いてあってその上に無作為に直線の棒を投げかけたとき、その棒が円を切る長さが三角形の一辺より長くなる確率を求めよ。」

この答えは、何に着目するかによって、1／2や1／3そして1／4と明らかに3種類の答えがでるのでパラドックスとされたものです。

しかし、現在ではこの問題の答えは分かっています。正解は1／2でした。回転、尺度、並進の3種理の変数変換に対して普遍である条件から導き出されます。1909年にボレルにより解決されています。





現在はラプラスの確率に関して提起された全てのパラドックスが解決されています。

フォン・ミーゼスの定義

ラプラスの定義では不満があるとして、フォン・ミーゼスは相対頻度の極限値を確率の定義としました。

「意図する実験がN回行われるものとし、事象Aが Na 回起こったとする。このとき、事象Aが起こる確率は比率 Na

／NでNを無限大にしたときの極限値とする。」

この定義が信頼性工学など、多くの工学分野で暗黙の内に使われている定義です。

この定義の確率でもコルモゴロフの公理は満たしますから数学の確率理論は使えます。

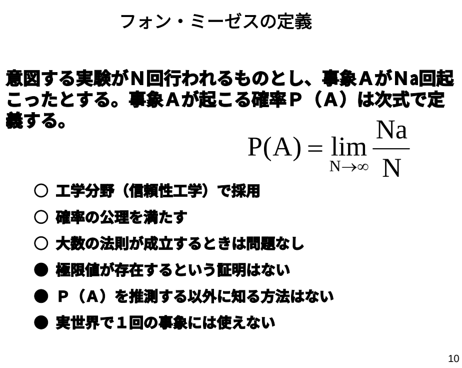
この定義で標本数が多いときは概ね問題がなかったのです。大量生産部品の品質管理手法に応用されています。数学的には大数の法則が成立するぐらい標本数が多いときは使えるということです。

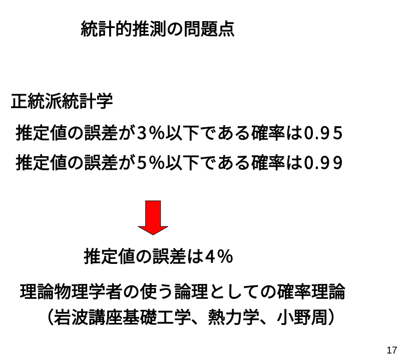
実は多すぎても問題があるのですが、このことは世の中にあまり知られていません。後で、「中国の皇帝の身長」という寓話を紹介します。

しかし、実世界の事象でこのような極限値が存在するとは言えません。また、フォン・ミーゼス自身が言っているのですが、一回しか起こりえないような事象には使えないのです。例えば、「第三次世界大戦が勃発する」確率などということには使えないのです。しかし、良く考えてみれば、世の中の事象はすべて1回限りと言えるかもしれません。

結局、P(A)の値も推測するしか手段がありません。すると推測には確かさの程度が問題になってきます。そこで信頼水準を置いて客観的に推定するということが行われています。

しかし、この方式が矛盾を含むものなのです。このことが、私がそもそも確率の応用がおかしいのではないかと考えたきっかけです。





頻度概念の確率では何故駄目かの実例はいくらでもあるのですが、ここでは一つ紹介いたします。それは正統派統計学といわれている現在の統計的推測のやり方です。

データを得て推定値の誤差を表現する場合、例えば、「誤差は3％以下であると95％の確かさで言える」という表現になってしまうのです。この表現だけでもすっきりしていませんが、同じデータを得て、「誤差は5％以下であると99％の確かさで言える」という表現も出来てしまうのです。

「確率3％以下であると95％信頼水準でいえる」という

表現の場合もあります。

　問題は信頼水準の取り方が恣意的であるということです。慣習的に６０％、９０％、９５％、９９％の数字を見かけます。場合によっては信頼水準の取り方によって結論が逆転します。

　意図的に結論が導けるから「その方が良いのだ」と言う人もいないわけではありませんが。

　もし、頻度概念確率を捨てて、確信の度合いの概念にすれば、このような問題の場合は「推定値の誤差は４％以下である」といったすっきりした表現が出来ます。

コルモゴロフの定義

　数学者のコルモゴロフは頻度概念とラプラス流との論争を離れ、次の公理を満たす P(A)を確率としました。

１）P(A)は非負数である。

２）確実な事象Ａの確率は１である。

３）もし、事象Ａと事象Ｂが相互背反ならば、ＡかＢが起こる確率 P(A+B)はＡの起こる確率 P(A)とＢの起こる確率 P(B)の和である。

つまり、確率は０から１の数値で表し、加法が成立するものとしたわけです。相互背反というのは同時に両方起こることはないということです。

　数学者は測度論に基づき、この確率の公理をゆるぎないものにしています。確率論はこの公理から演繹的に導かれる定理を集めたものですから論理的に正しいのです。

　ところが確率 P(A)は抽象的な数であって、実世界で何を意味しているのか、数学では何も言っていないのです。逆に、この公理を満たすものであったら、P(A)が何を意味するものであろうとも確率論の定理がすべて成立するということです。

サベイジの定義

　さて一番新しいのがサベイジの定義です。

これは「確率 P(A)とは命題Ａの真偽に関する確信の度合に割り当てた数値である」とするものです。命題とは真偽を明らかにすることを目的とした論述です。英語では proposition といいます。Preposition は前置詞ですから間違えないでください。

　例えば命題Ａとして「明日また太陽は昇る」ですとこの命題が眞であることは極めて強い確信を持てますから１とするでしょう。

　「１個のサイコロを振って偶数がでる。」はどうでしょうか。多くの方が確信の度合いは０．５であるということに納得されるでしょう。確信の度合いを確率の公理を満たすように決めることが出来ますので、確率の解釈として確信の度合いを

取ることも出来るわけです。

確かさの度合い（degree of belief）は明らかに人間の心の中で抱く数値であって、事象の属性ではありません。

サイコロの大きさや質量のような物理的属性でもありません。

コルモゴロフの定義

事象Ａの確率はこの事象に割り当てた数値Ｐ(Ａ)で次の３つの制約式を満たすものである。

１）Ｐ(Ａ)は正。 $P(A) \ge 0$

２）確実な事象の確率は１。 $P(S) = 1$

３）もしＡ、Ｂが相互背反ならば、

$$P(A+B) = P(A) + P(B)$$

数学者は測度論に基づいた公理から確率論を展開している。

Ｐ(Ａ)は抽象数で実世界でどのように使われるか関心がない。

11

サベイジの定義

Ｐ(Ａ)は命題の真偽に関する確信の度合いに割り当てた数値。

・情報が無いときは等確率の原理で割り当てる。

・データを見た後ではベイズの定理により計算する。

○ 確率の公理を満たすようにＰ(Ａ)をセット出来る。

○ この方法は意志決定の方法と合致している。

ベイズの定理：事後分布は
事前分布とデータの尤度の積に比例する。

12

確かさは物理量でなく心理量であると私は言っております。ベイズ統計学で用いられているのはこの確率です。通常事前確率を仮定し、データを見た後の事後確率をベイズの定理で求めるという方式を取ります。

本屋に行って確率の本を手に取りますと、実用的なもので頻度概念の確率とコルモゴロフの公理で終始しているものが多いようです。

数学的な確率の本はコルモゴロフの公理だけで済ましているものが多いのは当然でしょう。

ベイズ流と断ったものは、確率を確信の度合いの意味で解釈しています。

教科書的な本ですといろいろな定義があることを紹介しているものもあります。

しかし、これらの定義を並べて書いてあるだけでどれがどうだと書いてないのです。

確率は場合によって使い分けるべきものなのでしょうか。大学の先生は幾つもあるぞと学生に向かって知識を開陳すれば済むかもしれませんが、仕事に使う実務者としては一つでないと困るのです。

数学の確率は公理に基づく抽象的なものですから、問題はありません。むしろ色々な場合に共通に成立する方が学問的価値は高いでしょう。

実世界の確率はその意味をどのように捉えるべきかが関心事項なのです。

頻度概念確率はしばしば矛盾を生じたり不合理であったりします。これまで、確率とはそのようなものであると理解されて済まされてきたのです。

3. 確率の歴史

ワシントン大学の物理学の教授であった、E・T・ジェインズは長年の研究成果をまとめ「確率理論：科学の論理」という本の原稿を残しました。この原稿を数年掛けて体裁よくまとめて出版したのは、ジェインズの弟子のラリー・ブレットホーストです。

ここで、この本に書かれている確率の歴史を要約して紹介いたします。

(1) ベルヌーイ、ラプラスらが正しく確率理論を進めていた。

(2) この数理を良く理解できなかった生物学者、統計学者、哲学者の一派が。一見取り組み易い頻度概念の確率理論を打ち立ててしまった。しかし、これは願望投影錯誤(Mind Projection Fallacy) によるものである。

(3) ベン、フィッシャー、ネイマン、ピアソン、クラメール達はベイズ流の考え方を半ば感情的に攻撃し、正統派と呼ばれるようになった。(フィッシャーは自分の最後の論文では非を認めているとのこと)

(4) 正統派理論は適用できる範囲が極めて限られているにも係わらず、全てに応用しようとするのであちこちで矛盾を起こしている。一貫性の無い、その場しのぎの (ad hoc) 工夫で対処している。

(5) 論理としての確率理論は、ジェフリー、コックスの考え方に基づき、合理性と首尾一貫性の基本的な要求(デシダレータ)だけから論理的に展開するものである。

(6) 論理としての確率理論は、原則的に有限集合を基礎にするので実世界に応用してもパラドックスは生じない。あらゆる分野に適用できる。ラプラスの確率の定義も導ける。現代物理学の確率の考えにも合致している。

(7) 確率は推定するものでなく、情報に基づき割り当てるものである。どのように割り当てるかの原理として、最大エントロピー原理など数種類が見つけられている。

要約するとこの本が主張していることはこのようなことなのですが、頻度概念確率は不合理で矛盾を孕むものであるから捨てるべきだと言っています。

4. 拡張論理の確率理論

この拡張論理の確率が実世界の確率がどうあるべきかの回答を出しているのです。

拡張論理の確率理論がどういう位置づけにあるか、時間軸を横軸にした図にして見ました。

最初にラプラスの確率がありました。その次に頻度概念派が出てきて、ラプラスの確率を古典確率に追いやってしまい

ました。

頻度概念派で有名な学者はR．Aフィッシャーやクラメールです。

その次に、確信度を確率とする確信概念派が出て、頻度概念派と激しい論争を繰り広げました。確信概念派ではデ・フィネッティやD．V．リンドレーがいます。

数学者はその論争には直接参加せずに公理に基づく確率理論を打ち立てました。

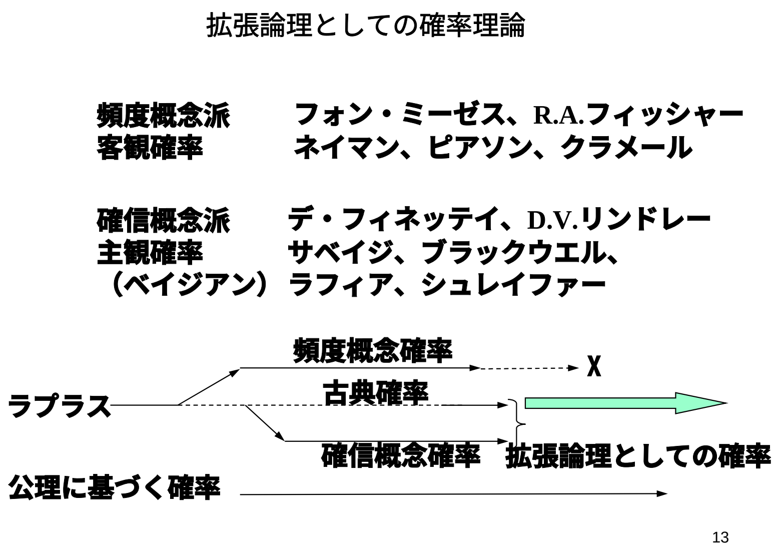
拡張論理の確率理論は、ラプラス流確率に基礎を与えるものであり、確信度を確率とするものに他ならないのです。頻度概念確率をきっぱり否定しています。

抽象概念としての公理に基づく確率理論には何も脅威を与えていませんが、実世界の確率は拡張論理の確率理論だけなのです。拡張論理の確率はラプラスの確率に基礎を与えたものですから物理学で使う確率とも整合が取れています。

また、通信理論を確立したシャノンが基礎にしたことは拡張論理の確率理論の基礎と同じことなので、現代の発達した情報理論とも整合が取れています。

拡張論理の確率理論に基づいた確率の決め方の一般規則はシャノンが導いた情報エントロピーを最大にする確率を求めることなのです。換言すると確率の決め方は情報エントロピー最大の原理によることなのです。

ジェインズは頻度概念確率が科学の進歩を阻害したとまで厳しく弾劾しています。頻度概念確率は完全に捨て去った方が良いのです。図では×印をつけておきました。



シャノンの通信理論を基にした情報分野が飛躍的に発展し

ているのに反し、信頼性工学に進歩がないのもこのためかもしれません。

拡張論理の確率理論がどのようなものか簡単に紹介します。

一般に確かさとは主観的なものであると思われています。何かが確かなことであるかどうかは、人によってその事に関する情報が異なるから、人によって確かさが違うのです。この意味で確かさは主観的であるというならば正しいでしょう。

しかし、情報に基づき確かさをどのように数値表現するのかの方法が決まっていれば、むしろ客観的な確かさと言えるでしょう。

簡単な例で、「サイコロ１個を振って６の目が出る」ということの確かさを考えてみます。

サイコロは６面体で１から６の目が出る可能性がどれも同じであると考えられますから等確率原理により数値を割り当てることができます。確率は１／６であると決めることは誰もが納得できるのではないでしょうか。

サイコロ振りのように単純な場合だけでなく、複雑な場合にも使える方法は何だろうかを突き詰めたものが拡張論理の確率理論です。

基礎的な事項として次の三つを掲げます。

・合理的であること。

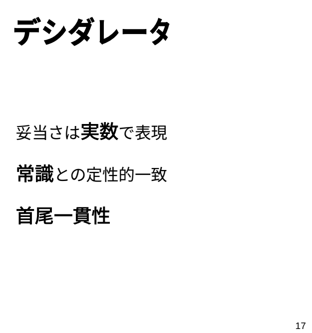
・常識との一致。

・首尾一貫していること。

これだけなのです。もしかすると常識との一致が難しいと思われる方がおられるかもしれません。世の中の常識とは変わり得るものだからです。

しかし、幸いなことにここでは、確かなことを確率１で表し、絶対にありえない事を確率０で表すという方向性を決めるだけにしか使われていません。全く逆にしても成立してしまうからです。

首尾一貫しているということは、矛盾がないということです。



ジェインズは論理を進めるにあたってこの三つを基礎にすることにし、これらをデシダレータと呼びました。デシダレータとはラテン語で不可欠なものという意味です。三つならべているので単数形のデシダレータムでなくデシダレータです。

世の中の事象はすべて不確かであってその確かさに程度があるだけなのだと考えるのです。

命題が眞である妥当性の程度を表すために数値を与えるのです。その与え方の原則として、デシダレータを採用するというものです。

論理学の論理の進め方はAならばBであるという状況下で言えることは二つだけです。つまりAだったのでBであるということと、Bでなかったので A ではない。この論法だけで組み立てた理論は如何に複雑であっても有無を言わさず正しいものです。

現実の世界では、AならばBであるという状況下でAでなかったという場合も多いのです。この場合演繹的な論理では何も言えないのですが、少し言えることがあるのです。

つまり、AならばBであるという状況下でAでなかったとき、Bでない可能性が強くなったとは言えるのです。

このような関係も論理学と同じようにブール代数が成立すると考えて推論を進めるところに拡張論理の拡張の意味するところです。

この論理を進めますと積の定理と和の定理が導けるのです。そして、違いが認識できないものは同じであるという無差別の原理と、確実な事象には妥当性を1で表すということにします。すると、導かれた結果は結局ラプラスの確率定義に一致するものであったということなのです。

これは確信度を数値で表した確率でもあるのです。

5. 実世界は有限

拡張論理の確率理論は本質的に有限集合を背景に持てば良いとジェインズは言っています。実世界は有限だからです。

この点は純粋数学の公理に基づく確率理論とは違います。コルモゴロフの確率は測度論により磐石の基礎を持っていますが、その背景は無限集合です。コルモゴロフの確率はあくまでも数学という抽象の世界です。

現実世界は有限であることを実感してみましょう。また、現実世界は有限であるだけでなく、量子的に不連続です。

一番大きなものを一番小さいもので図ってどのくらいの数字になるかで有限の大きさの程度が分かります。

有名な歌舞伎のセリフに「石川や浜の真砂は尽きるとも世に盗人の種は尽きまじ」というのがあります。石川五右衛門が釜茹での刑に処せられたときの辞世の句だそうです。

海岸の砂浜の砂の数は有限であり数えられることを知っていたのでしょうか。

余談ですが、日本で一番広い砂浜を持つのは千葉県の九十九里浜でしょうか。さて九十九里浜の砂粒の数は何個ぐらいかと言う問題に答えられますか。

このような問題はフェルミの問題と言われています。持てる知識を総動員し、頭を使ってとにかく概数を推測するわけです。ビル・ゲイツがマイクロソフトに入る入社試験にはこのような問題を出すそうです。

有限の話に戻りまして、まず長さで考えてみます。この世の中で一番小さいものは素粒子で10の-15乗mです。宇宙の大きさは137億光年ですから有限です。宇宙の大きさ素粒子の大きさで勘定しても10の200乗にもならないでしょう。まだ実在が確認されていませんが、「超ひも理論」の「超ひも」は10の-35乗mぐらいであると言われています。超ひもで図っても20桁大きな数字になるだけです。

次に質量で見てみます。現実世界で考えられる一番大きな質量は宇宙全体の質量でしょう。

太陽の質量は2×10の33乗gです。銀河は2×10の11乗個の太陽程度の大きさの星を持っています。宇宙には我々の太陽系が含まれて居るような銀河がやはり2×10の11乗個あるそうです。これらを全部掛け合わせると宇宙全体の質量が計算できます。ダークマターがその何倍かあるらしいということですのでそれらも勘定に入れても有限であることに変わりはないでしょう。

水素原子の質量は1.6×10の-24乗gです。水素原子の質量を単位としても、まだ単位が大きすぎるということであれば、電子の質量を単位として図っても良いでしょう。

要するに質量があると考えられている素粒子の中で一番小さい質量を単位にとれば良いのです。それで現実世界の量は有限であることが実感できます。

ここで注意すべきなのは、組み合わせの数は有限であっても容易に実世界で存在し得る数を超えて大きくなるということです。抽象の世界に入ってしまうほどです。

トランプの52枚のカードを並べる並べ方で、カードをいくら切っても、人類がまだ並べたことのない順序はいくらでもあることになります。

実世界で無限大はないことは、逆に無限小も実世界にはないということです。このことから連続の概念も抽象的概念であることが分かります。

大学の数学の授業で、ε、δ論法が出てきて悩まされた覚えがあります。このような微小の概念はまさに抽象の世界です。

ラプラスの確率に関連するパラドックス問題は、殆どすべ

て無限の概念から来ています。従って、実世界の問題ではパラドックスは起こりえないと考えて良いのです。

果たして時間は連続に流れているのだろうか。かなり哲学的な問いかけです。

現代物理学は時間に関しても最小の時間間隔があると考えることにより、物理現象がうまく説明できるとしています。これはプランク時間と呼ばれています。

6.ラプラスの連続則

サイコロは最初から6面体であるという状況が分かっています。一般的には最初は何も判っていないという状況が殆どです。

不確かな状況から一つ二つと情報を得ることによってだんだん確かになっていきます。この状況を最も簡単なモデルに対して、データに基づきどのように確かさを変化させるかを示したのがラプラスです。

モデルとしてはA社が開発した分離ボルトであったとしましょう。分離ボルトは火薬で作動させるもので、結果は成功か失敗かが明確にどちらかにでます。

A社についてはどんな会社なのか、信頼のおける会社なのか、全く分からないのです。A社から調達する新開発の分離ボルトは確実に作動するでしょうか。

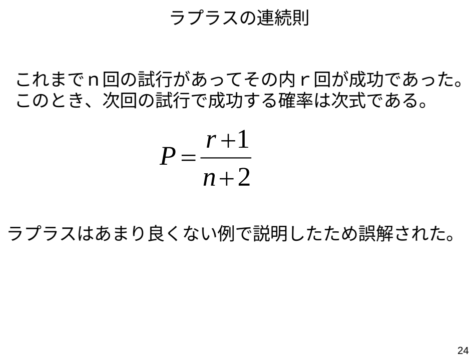
このような場合、通常何個か買って試してみます。

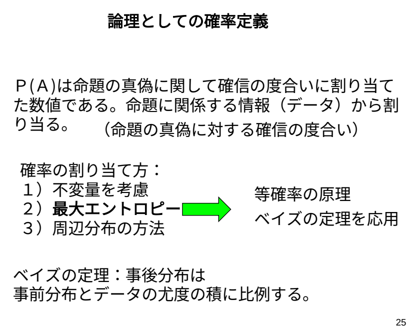
始めてこの分離ボルトを入手した時点では、動作するか否かの二者択一ですから等確率の原理で確率は0.5です。

1個試して成功すると、大丈夫そうだなと言う感触を得ます。5，6個試してみて全部成功すれば、絶対大丈夫と言えるでしょうか。

n個試験してr個成功したのを見た後では、確率は(r+1)/(n+2)になるというものです。この式はラプラスの連続則と言う名前が付けられています。

nに色々な数字を当てはめてみると実感と合うと思います。





最初から失敗無く8個の成功を見て、やっと確かさは0.

９であると言えるのです。

ラプラスはこの式を紹介するときに、あまり良くない例を使ったため、逆に誤解を受けたようです。

ラプラスの連続則は結果の式は易しいのですが、誘導はかなり込み入っていますので、ここでは紹介を省略します。

復習しますと、拡張論理の拡張を省略して、論理としての確率は「命題の真偽に関し確信の度合いに割り当てた数値である。」ということです。そして数値の割り当て方は命題に関係する情報から割り当てるものです。

数値の割り当て方は、最大エントロピー原理によりますが、その他に不変量を考慮するもの、他があります。

事前分布として等確率の原理を使い、データを得た後は事前分布とベイズの定理によるというのが原則的に行う方法です。

拡張論理としての確率理論

ここで拡張論理としての確率理論の位置づけをまとめておきます。

まず論理的推論の帰結としてラプラス流の確率が導かれます。

その推論の基礎にしたのは三つのデシダレータでした。拡張した論理にもブール代数を適用することにより、積の規則と和の規則が導かれます。そして、無差別の原理から、絶対的に強い確信を1として、ラプラス流の確率定義が誘導できるのです。

ジェインズはC．E．シャノンにも影響を与えていて、シャノンが通信理論の基礎としたシャノンの条件はデシダレータを別の側面から見たものに他ならないのです。

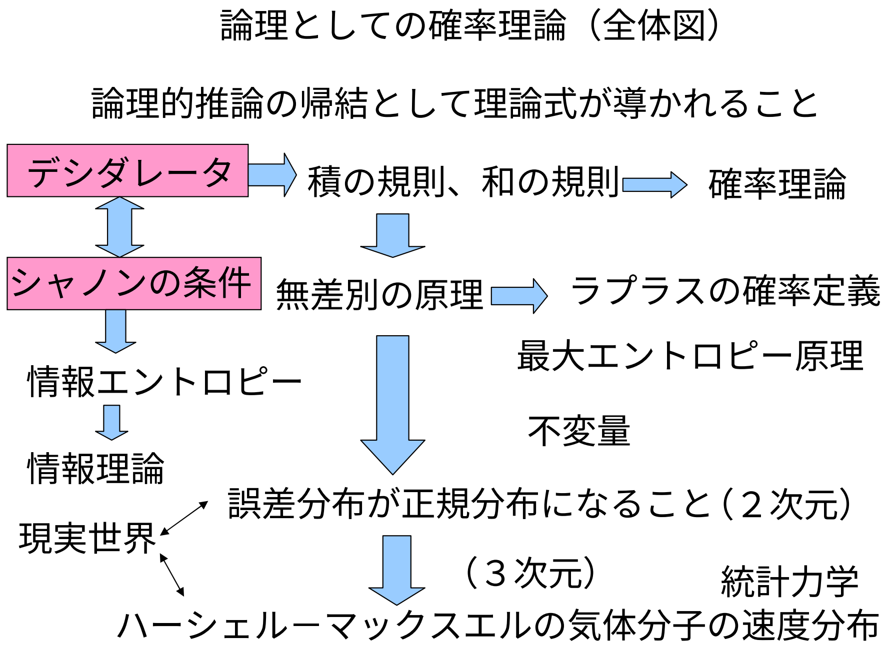
そして、確かさを数値で表すには、「シャノンの情報エントロピーを最大にするような確率を求める」ことが首尾一貫した合意理的な方法であることになるのです。

ただし、複雑な実社会では情報エントロピーの計算は非常に難しくなります。

一方、不変量を考慮することから確率を決めることができます。この例として既にバートランドのパラドックスの解決を紹介しています。しかし、もっと重要な例があります。

有名なガウスの誤差分布は正規分布でもありますが、この分布が論理的な考察からだけで誘導できるのです。

物理学では同じような考察だけから導かれたものに、ハーシェル・マックスウエルの気体分子の速度分布があります。



26

7．関数方程式

確率統計の分野で最も重要な分布が正規分布でしょう。この式が論理的な考察だけから導かれるものであることを述べましたが、その考察で重要なのが関数方程式です。

私達の数学の授業では出てきませんので、関数方程式がどのようなものであるか説明します。

関数方程式とは未知関数F(X)を含む式です。代表的な例はコーシーの方程式です。これはF(X+Y)=F(X)+F(Y)です。

関数方程式はその関数の持っている性質を表しています。

正規分布

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{\frac{-(x-\mu)^2}{2\sigma^2}\right\},\qquad -\infty<x<\infty,\quad \sigma^2>0$$

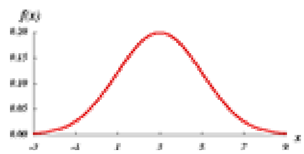


図　正規分布 N(3, 2^2) の概形

27

関数方程式

関数方程式：未知関数 F(X) を含む方程式

F(X+Y)=F(X)+F(Y)　：コーシーの方程式

関数方程式は、その関数の持っている性質を表している

関数方程式における興味の中心は、その性質を満たす関数をすべて求めたい

その他の関数方程式

（1）全ての実数 X、Y に対して、F(X＋Y)＝F(X)F(Y)
（2）全ての正数 X、Y に対して、F(XY)＝F(X)＋F(Y)
（3）全ての正数 X、Y に対して、F(XY)＝F(X)F(Y)

28

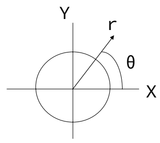
それではガウスの誤差分布がどのように導出されるのかを

見てみましょう。

今、X，Y座標の中心に向かって矢を射ると想定しましょう。この矢が当る場所の確率密度が求めるもので、これをρ(x,y)とします。矢の散らばり方は、X方向Y方向で違う理由がありませんので,f(x)f(y)と表せるものでなければなりません。

直交座標でなく極座標で表現したとすると、ρ(x,y)は g(r,θ)に等しいはずです。矢の散らばり方は角度によって違いがないはずですから g(r,θ)は g（r）と書けなければならないのです。

ガウスの誤差分布



的の中心に向かって矢を射る

幾何学的に、直感的に、要求される均等性の二つの条件

直角方向の誤差の確率は独立であるべき

$$\rho(x,y)dxdy = f(x)f(y)dxdy$$

極座標でも $\rho(x,y)dxdy = g(r,\theta)rdrd\theta$

θ に独立であるべきだから、$g(r,\theta) = g(r)$

30

ガウスの誤差分布

従って、 $f(x)f(y) = g(\sqrt{x^2+y^2})$

$$\log\frac{f(x)}{f(0)} + \log\frac{f(y)}{f(0)} = \log\frac{f(\sqrt{x^2+y^2})}{f(0)}$$

この関数方程式を満たす関数は $\log\frac{f(x)}{f(0)} = ax^2$

従って、 $f(x) = f(0)e^{ax^2}$

正規化条件で係数を決めると、

$$f(x) = \sqrt{\frac{\alpha}{\pi}}e^{-\alpha x^2}, \qquad \alpha > 0$$

31

結局、図にあるようなｘ、とｙをについての関数方程式が得られて、これを満たす関数は指数関数で正規分布と言われる式でなければならないことが結論できるのです。

正規化条件で未知パラメータが一つ決まります。正規化条件というのは。確率密度を全区間に渡り積分した値が１であるというものです。

もう一つのパラメータは分布の散らばりの程度に対応するものですが、これは事象の性質に依存します。弓矢であれば射手の上手さの程度でしょう。

8．中国の皇帝の身長

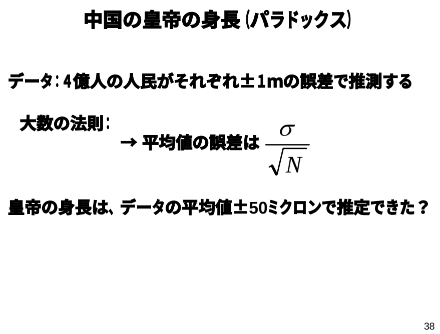
この話もジェインズの本からの抜粋です。数字は多少変えてあります。面白いパラドックスですが、確率理論の応用に重要な注意を促すものです。

ある時代の中国の皇帝が４億人の民を治めていました。この民は自分たちを納めている皇帝の身長がどのくらいか知りたいと思いました。皇帝の体に物差しを当てることなど畏れ多くて誰にも出来ません。ある人は、１ｍ６０ｃｍぐらいだ、別の人は１ｍ８０ｃｍぐらいだと言いますが誰も自信を持って答えることは出来ません。

そこで、プラスマイナス１ｍの誤差を許したら誰もが自信を持って推定することが出来ました。中国人民全員から１ｍの誤差付の推定値を得たわけです。４億のデータを平均し、

誤差を計算すると、平均値の誤差は何と５０ミクロンの小さ
さになります。中国人民は皇帝の身長を５０ミクロンの誤差
で知ることができたのでしょうか。

現代の統計学で最も重要な定理の一つに中央極限定理があ
ります。この定理は誤差σを伴う測定値がｎ個あった時に、
その平均値の誤差はσ／√ｎであることを教えています。



このパラドックスはどこに間違いが潜んでいるのかは次の
例を考えると明らかになります。

ある物の長さを出来るだけ精度よく計るとします。一人の
人が精度１／１００ｍｍで測定できるマイクロメータを使っ
て１回測定したデ値がＡｍｍであったとします。同じものを
１００人の人が精度１／１０ｍｍのノギスを使って測定し、
得られた１００個のデータの平均をＢｍｍとします。Ｂの誤
差は同じく、マイクロメータの精度と同じく１／１００ｍｍ
でしょうか。

ノギスを使う計測者が１０００人であったら、マイクロメ
ータの測定精度をしのぐのでしょうか。

おわりに

世の中に使われている確率の意味がいくつかあることをお
話しました。その中で、現在多くの工学分野で使われている
頻度概念確率は不適切なものであることを強調いたしました。

私達が採用すべきは論理としての確率理論であり、合理的
で首尾一貫した論理的帰結としての確率理論なのです。

これは現代の物理学で採用されているものでもあり、近年
目覚しい発達を成し遂げた通信分野の基礎理論にも採用され
ているものです。

現在使われている工学分野でも頻度概念確率を捨て去りこ
れまでの方式を見直すことで新たな進展が期待できるものと
思われます。

（平成20年7月2日）