

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2003-281116

(P2003-281116A)

(43)公開日 平成15年10月3日(2003.10.3)

(51)Int.Cl.⁷

G 0 6 F 17/18

識別記号

F I

G 0 6 F 17/18

タームコード*(参考)

Z 5 B 0 5 6

審査請求 未請求 請求項の数6 O L (全 7 頁)

(21)出願番号 特願2002-83110(P2002-83110)

(22)出願日 平成14年3月25日(2002.3.25)

(71)出願人 800000080

タマティールエルオー株式会社

東京都八王子市旭町9番1号 八王子スク

エアビル11階

(72)発明者 高橋 静昭

東京都新宿区西新宿1-24-2 工学院大

学内

(74)代理人 100094053

弁理士 佐藤 隆久

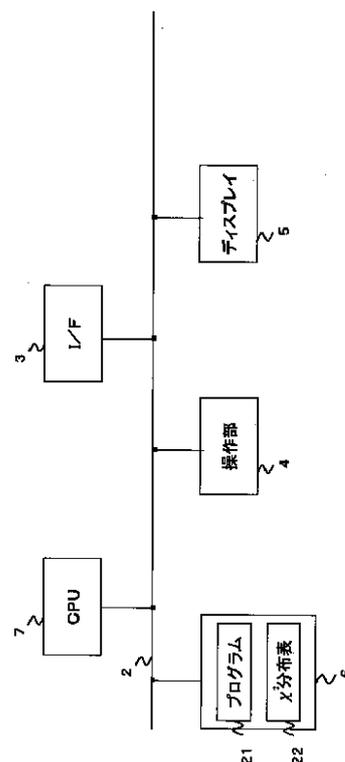
Fターム(参考) 5B056 BB64 BB72

(54)【発明の名称】 データ処理装置、その方法およびそのプログラム

(57)【要約】

【課題】 観測条件を観測者が自由に決定できる場合に、従来に比べて信頼性が高く、しかも、簡単な構成で、短時間に、観測条件が所定の工程に影響を与えているか否かを判定できるデータ処理装置を提供する。

【解決手段】 第1の条件の基に行われたN回の工程で $E_1 \sim E_s$ の事象がそれぞれ $f_1 \sim f_s$ 回生じ、第2の条件の基に行われたN回の前記工程で前記 $E_1 \sim E_s$ の事象がそれぞれ $g_1 \sim g_s$ 回生じた場合に、CPU 7 が、データ $f_1 \sim f_s$ および $g_1 \sim g_s$ を用いて所定の演算を行ってデータDを算出し、指定された有意水準と自由度 $(S - 1)$ との χ^2 分布とで規定される $S-1$ () と前記データDとを比較し、データDが $S-1$ () より大きい場合に、前記第1の条件および前記第2の条件が、前記工程で生じる前記事象に影響を与えていると判定する。



【特許請求の範囲】

【請求項1】第1の条件の基に行われたN回の工程でE₁ ~ E_sの事象がそれぞれf₁ ~ f_s回生し、第2の条件の基に行われたN回の前記工程で前記E₁ ~ E_sの事象がそれぞれg₁ ~ g_s回生した場合に、データf₁ ~ f_sおよびg₁ ~ g_sを入力する入力手段と、前記入力されたデータf₁ ~ f_sおよびg₁ ~ g_sを用いて、下記式(1)で規定される演算を行ってデータD*

$$D = (\sqrt{2f_1} - \sqrt{2g_1})^2 + (\sqrt{2f_2} - \sqrt{2g_2})^2 + \dots + (\sqrt{2f_s} - \sqrt{2g_s})^2$$

・・・(1)

【請求項2】前記第2の条件は、前記第1の条件ではないという条件である請求項1に記載のデータ処理装置。

【請求項3】前記自由度(S-1)と前記有意水準とから前記s-1()を一意に決定する²分布表データを記憶する記憶手段をさらに有し、前記比較手段は、有意水準と自由度(S-1)とを用いて、前記分布表データを検索して前記s-1()を

【請求項4】s-1()を算出するs-1()算出手段をさらに有し、前記比較手段は、前記算出されたs-1()を用いて前記比較を行う請求項1に記載のデータ処理装置。

【請求項5】第1の条件の基に行われたN回の工程でE₁ ~ E_sの事象がそれぞれf₁ ~ f_s回生し、第2の条件の基に行われたN回の前記工程で前記E₁ ~ E_sの事

$$D = (\sqrt{2f_1} - \sqrt{2g_1})^2 + (\sqrt{2f_2} - \sqrt{2g_2})^2 + \dots + (\sqrt{2f_s} - \sqrt{2g_s})^2$$

・・・(2)

【請求項6】データ処理装置が行う処理内容が記述され、前記データ処理装置が機械的に読み取り可能なプログラムであって、

第1の条件の基に行われたN回の工程でE₁ ~ E_sの事象がそれぞれf₁ ~ f_s回生し、第2の条件の基に行われたN回の前記工程で前記E₁ ~ E_sの事象がそれぞれg₁ ~ g_s回生した場合に、データf₁ ~ f_sおよびg₁ ~ g_sを入力する手順と、前記入力されたデータf₁ ~ f_sおよびg₁ ~ g_sを用いて、下記式(3)で規定される演算を行ってデータD

$$D = (\sqrt{2f_1} - \sqrt{2g_1})^2 + (\sqrt{2f_2} - \sqrt{2g_2})^2 + \dots + (\sqrt{2f_s} - \sqrt{2g_s})^2$$

・・・(3)

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、2つの確率分布の適合度を検定するデータ処理装置、その方法およびその

*を算出する演算手段と、

指定された有意水準と自由度(S-1)との²分布とで規定されるs-1()と前記データDとを比較する比較手段と、

前記比較の結果、データDがs-1()より大きい場合に、前記第1の条件および前記第2の条件が、前記工程で生じる前記事象に影響を与えていると判定する判定手段とを有するデータ処理装置。

【数1】

象がそれぞれg₁ ~ g_s回生した場合に、データf₁ ~ f_sおよびg₁ ~ g_sを入力するステップと、

前記入力されたデータf₁ ~ f_sおよびg₁ ~ g_sを用いて、下記式(2)で規定される演算を行ってデータDを算出するステップと、

20 指定させた有意水準と自由度(S-1)との²分布とで規定されるs-1()と前記データDとを比較するステップと、

前記比較の結果、データDがs-1()より大きい場合に、前記第1の条件および前記第2の条件が、前記工程で生じる前記事象に影響を与えていると判定するステップとを有するデータ処理方法。

【数2】

を算出する手順と、

指定された有意水準と自由度(S-1)との²分布とで規定されるs-1()と前記データDとを比較する手順と、

前記比較の結果、データDがs-1()より大きい場合に、前記第1の条件および前記第2の条件が、前記工程で生じる前記事象に影響を与えていると判定する手順とを有するプログラム。

【数3】

プログラムに関する。

【0002】

【従来の技術】S通りの事象(出来事)E₁, E₂, ..., E_sのうちどれが起こるかを観測する場合

に、それを観測するための条件が、Aという観測条件の場合とAでない(A₂)という観測条件の2通りの場合がある。この観測条件の違いが、上記S通りの事象の起き方に何らかの影響を及ぼすか否かを判断したい場合がある。従来では、この判断を以下に行っている。第1の手法では、観測条件が、Aであるか否かを確率的な事象であると考え、観測条件Aと事象E₁, E₂, . . . , E_sとが独立しているか否かを、周辺分布を観測して²検定を行う。また、第2の手法では、2項分布に従う変数Xに対して、SIN⁻¹(X^{1/2})が正規分布に従うことを利用して²検定を行う。

【0003】

【発明が解決しようとする課題】しかしながら、上述した第1の手法では、一般に観測条件Aか否かが確率的な要因ではなく、観測者が自由に決定できる場合が多く、この場合に、両者の観測条件としてAとA₂のそれぞれの観測回数が定数であるにもかかわらず、それらを確率変数として扱ってしまうため、判定結果の信頼性が低いという問題がある。また、上述した第2の手法では、関数「SIN⁻¹(X^{1/2})」の演算処理を伴い、コンピュータで実行する場合に、プログラムが複雑になると共に、処理時間が長いという問題がある。

【0004】本発明はかかる事情に鑑みてなされたもの*

$$D = (\sqrt{2f_1} - \sqrt{2g_1})^2 + (\sqrt{2f_2} - \sqrt{2g_2})^2 + \dots + (\sqrt{2f_s} - \sqrt{2g_s})^2$$

・・・(4)

【0007】第1の発明のデータ処理装置の作用は以下のようになる。入力手段、データf₁ ~ f_sおよびg₁ ~ g_sを入力する。次に、演算手段、前記入力されたデータf₁ ~ f_sおよびg₁ ~ g_sを用いて、上記式(4)で規定される演算を行ってデータDを算出する。次に、比較手段が、指定された有意水準と自由度(S-1)との²分布とで規定される_{s-1}()と前記データDとを比較する。次に、判定手段が、前記比較の結果、データDが_{s-1}()より大きい場合に、前記第1の条件および前記第2の条件が、前記工程で生じる前記事象に影響を与えていると判定する。

【0008】第1の発明のデータ処理装置は、好ましくは、前記第2の条件は、前記第1の条件ではないという条件である。また、第1の発明のデータ処理装置は、好ましくは、前記自由度(S-1)と前記有意水準とから前記_{s-1}()を一意に決定する²分布表データを記憶する記憶手段をさらに有し、前記比較手段は、有意水準と自由度(S-1)とを用いて、前記分布表データを検索して前記_{s-1}()を特定する。また、第1の発明のデータ処理装置は、好ましくは、_{s-1}

*であり、観測条件を観測者が自由に決定できる場合に、従来に比べて信頼性が高く、しかも、簡単な構成で、短時間に、観測条件が所定の工程に影響を与えているか否かを判定できるデータ処理装置、その方法およびそのプログラムを提供することを目的とする。

【0005】

【課題を解決するための手段】上記の目的を達成するため、第1の発明のデータ処理装置は、第1の条件の基に行われたN回の工程でE₁ ~ E_sの事象がそれぞれf₁ ~ f_s回生じ、第2の条件の基に行われたN回の前記工程で前記E₁ ~ E_sの事象がそれぞれg₁ ~ g_s回生じた場合に、データf₁ ~ f_sおよびg₁ ~ g_sを入力する入力手段と、前記入力されたデータf₁ ~ f_sおよびg₁ ~ g_sを用いて、下記式(4)で規定される演算を行ってデータDを算出する演算手段と、指定された有意水準と自由度(S-1)との²分布とで規定される_{s-1}()と前記データDとを比較する比較手段と、前記比較の結果、データDが_{s-1}()より大きい場合に、前記第1の条件および前記第2の条件が、前記工程で生じる前記事象に影響を与えていると判定する判定手段とを有する。

【0006】

【数4】

$$D = (\sqrt{2f_1} - \sqrt{2g_1})^2 + (\sqrt{2f_2} - \sqrt{2g_2})^2 + \dots + (\sqrt{2f_s} - \sqrt{2g_s})^2$$

・・・(4)

()を算出する_{s-1}()算出手段をさらに有し、前記比較手段は、前記算出された_{s-1}()を用いて前記比較を行う。

【0009】また、第2の発明のデータ処理方法は、第1の条件の基に行われたN回の工程でE₁ ~ E_sの事象がそれぞれf₁ ~ f_s回生じ、第2の条件の基に行われたN回の前記工程で前記E₁ ~ E_sの事象がそれぞれg₁ ~ g_s回生じた場合に、データf₁ ~ f_sおよびg₁ ~ g_sを入力するステップと、前記入力されたデータf₁ ~ f_sおよびg₁ ~ g_sを用いて、下記式(5)で規定される演算を行ってデータDを算出するステップと、指定された有意水準と自由度(S-1)との²分布とで規定される_{s-1}()と前記データDとを比較するステップと、前記比較の結果、データDが_{s-1}()より大きい場合に、前記第1の条件および前記第2の条件が、前記工程で生じる前記事象に影響を与えていると判定するステップとを有する。

【0010】

【数5】

$$D = \left(\sqrt{2f_1} - \sqrt{2g_1}\right)^2 + \left(\sqrt{2f_2} - \sqrt{2g_2}\right)^2 + \dots + \left(\sqrt{2f_s} - \sqrt{2g_s}\right)^2 \dots (5)$$

【0011】また、第3の発明のプログラムは、データ処理装置が行う処理内容が記述され、前記データ処理装置が機械的に読み取り可能なプログラムであって、第1の条件の基に行われたN回の工程でE₁ ~ E_sの事象がそれぞれf₁ ~ f_s回生じ、第2の条件の基に行われたN回の前記工程で前記E₁ ~ E_sの事象がそれぞれg₁ ~ g_s回生じた場合に、データf₁ ~ f_sおよびg₁ ~ g_sを入力する手順と、前記入力されたデータf₁ ~ f_sおよびg₁ ~ g_sを用いて、下記式(6)で規定され*

$$D = \left(\sqrt{2f_1} - \sqrt{2g_1}\right)^2 + \left(\sqrt{2f_2} - \sqrt{2g_2}\right)^2 + \dots + \left(\sqrt{2f_s} - \sqrt{2g_s}\right)^2 \dots (6)$$

【0013】

【発明の実施の形態】以下、本発明の実施形態に係わるコンピュータについて説明する。図1は、本実施形態のコンピュータ1の構成図である。図1に示すように、コンピュータ1は、例えば、インタフェース3、操作部4、ディスプレイ5、メモリ6およびCPU(Central Processing Unit)7を有し、これらがバス2を介して接続されている。ここで、インタフェース3が本発明の入力手段に対応し、メモリ6が本発明の記憶手段に対応している。

【0014】インタフェース3は、例えば、コンピュータ1の外部装置との間でデータ入出力を行う。操作部4は、キーボードやマウスなどの操作手段であり、ユーザによる操作に応じた操作信号を、バス2を介してCPU7に出力する。ディスプレイ5は、CPU7からの表示信号に応じた画面を表示する。

【0015】メモリ6は、プログラム21および²分布表データ22を記憶している。プログラム21は、後述するCPU7が行う処理の手順を記述したプログラムである。²分布表データ22は、例えば、図2に示すように、自由度(S-1)と、基準の確率(ある仮説が

$$D = \left(\sqrt{2f_1} - \sqrt{2g_1}\right)^2 + \left(\sqrt{2f_2} - \sqrt{2g_2}\right)^2 + \dots + \left(\sqrt{2f_s} - \sqrt{2g_s}\right)^2 \dots (7)$$

【0019】本実施形態では、例えば、ある製品を製造する工場で製品がその品質から1, 2, 3, 4の等級に分けられ、製造工程で、ある材料に特定の薬品を使用した場合と使用しなかった場合の2通りについて、製品の等級の個数を調べ、当該薬品の使用が製品の等級の分布に与える影響があるか無いかを考える。そして、薬品の

*の演算を行ってデータDを算出する手順と、指定された有意水準と自由度(S-1)との²分布とで規定される_{s-1}()と前記データDとを比較する手順と、前記比較の結果、データDが_{s-1}()より大きい場合に、前記第1の条件および前記第2の条件が、前記工程で生じる前記事象に影響を与えていると判定する手順とを有する。

【0012】

【数6】

正しいにもかかわらず正しくないと判断する確率)である有意水準とから_{s-1}()を特定する表データである。また、Sは、所定の工程で生じる事象の数を示している。

【0016】CPU7は、プログラム21を基に動作し、後述する処理を行う。図3に示すように、CPU7は、確率変数演算部11、比較部12および判定部13とを有する。なお、CPU7は、実際には、プログラム21を基に動作し、確率変数演算部11、比較部12および判定部13に相当する機能の処理を行う。

【0017】なお、確率変数演算部11は、図4(A)に示すように、第1の条件の基に行われたN回の工程でE₁ ~ E_sの事象がそれぞれf₁ ~ f_s回生じ、第2の条件の基に行われたN回の前記工程で前記E₁ ~ E_sの事象がそれぞれg₁ ~ g_s回生じた場合に、データf₁ ~ f_sおよびg₁ ~ g_sを用いて、下記式(7)で規定される演算を行って確率変数であるデータDを算出する。

【0018】

【数7】

使用の有無と各等級の個数とが、図4(B)に示すようになった場合には、データDは、下記式(8)のようになる。この場合には、S=4となる。

【0020】

【数8】

$$D = \left(\sqrt{2 \times 254} - \sqrt{2 \times 202} \right)^2 + \left(\sqrt{2 \times 590} - \sqrt{2 \times 595} \right)^2 + \left(\sqrt{2 \times 520} - \sqrt{2 \times 540} \right)^2 + \left(\sqrt{2 \times 136} - \sqrt{2 \times 183} \right)^2 \approx 8.39235$$

・・・ (8)

【0021】比較部12は、操作部4を介してユーザから入力された有意水準、あるいは、予め決められた有意水準と、自由度(S-1)とをキーとして、図2に示す²分布表データ22を参照して、 s_{-1} ()を得る。そして、比較部12は、当該得た s_{-1} ()と、確率変数演算部11が算出したデータDとを比較する。判定部13は、比較部12の比較の結果を基に、データDが s_{-1} ()より大きい場合には第1および第2の条件が工程で生じる事象に影響を与えていると判定し、そうでない場合には第1および第2の条件が工程で生じる事象に影響を与えていないと判定する。具体的には、判定部13は、図4(B)に示す例では、比較部12の比較の結果を基に、データDが s_{-1} ()より大きい場合には薬品の使用が製品の等級に影響を与えていると判定し、そうでない場合には薬品の使用が製品の等級に影響を与えていないと判定する。図4(B)に示す例では、上記式(8)に示すようにデータDが「約8.39235」となり、 s_{-1} (0.05)が「7.81473」となり、 $D > s_{-1}$ (0.05)であるため、薬品の使用が製品の等級に影響を与えていると判定される。

【0022】以下、図1に示すコンピュータ1の動作例を説明する。図5は、当該動作例を説明するためのフローチャートである。

ステップST1：ユーザが、第1の条件の基にN回の工程を行い、第2の条件の基にN回の工程を行い、第1の条件の工程で $E_1 \sim E_s$ の事象がそれぞれが何回生じたか、並びに第1の条件の工程で $E_1 \sim E_s$ の事象がそれぞれが何回生じたかを調べて、前述したデータ $f_1 \sim f_s$ および $g_1 \sim g_s$ を収集する。

【0023】ステップST2：ステップST1で収集されたデータ $f_1 \sim f_s$ および $g_1 \sim g_s$ が、インタフェース3あるいは操作部4を介してコンピュータ1に入力され、例えば、メモリ6に記憶される。

【0024】ステップST3：CPU7の確率変数演算部11が、ステップST2でメモリ6に記憶されたデータ $f_1 \sim f_s$ および $g_1 \sim g_s$ を用いて、上記式(7)に示す演算を行い、データDを算出する。

【0025】ステップST4：CPU7の比較部12が、操作部4を介してユーザから入力された有意水準、あるいは、予め決められた有意水準と、自由度(S-1)とをキーとして、図2に示す²分布表データ22を参照して、 s_{-1} ()を取得する。

【0026】ステップST5：比較部12が、当該得た s_{-1} ()と、ステップST3で確率変数演算部11

が算出したデータDとを比較する。

【0027】ステップST6：判定部13が、ステップST5の比較の結果を基に、データDが s_{-1} ()より大きい場合にはステップST7の処理に進み、そうでない場合にはステップST8の処理に進む。

【0028】ステップST7：判定部13が、第1および第2の条件が工程で生じる事象に影響を与えていると判定し、例えば当該判定結果を示す画面をディスプレイ5に表示する。

ステップST8：判定部13が、第1および第2の条件が工程で生じる事象に影響を与えていない判定し、例えば当該判定結果を示す画面をディスプレイ5に表示する。

【0029】以上説明したように、コンピュータ1によれば、第1の条件および第2の条件下である観測回数Nが定数である場合に、それを従来の第1の手法のように確率変数ではなく、定数として扱うため、信頼性の高い判定を行うことができる。また、コンピュータ1によれば、従来の第2の手法のように、関数「 $SIN^{-1}(X^{1/2})$ 」などの複雑な演算を行う必要がないため、プログラム21を簡単にでき、しかも処理時間を短縮できる。

【0030】本発明は上述した実施形態には限定されない。例えば、上述した実施形態では、図1に示すメモリ6に記憶された²分布表データ22を用いて s_{-1} ()を取得する場合を例示したが、本発明では、例えば、²分布表データ22を用いずに、CPU7が所定の演算を行って s_{-1} ()を算出してもよい。例えば、CPU7は、自由度 $n (= S - 1)$ が3の場合に、下記式(9)を満たす λ を算出する。

【0031】

【数9】

$$\int_{\lambda}^{\infty} \frac{X^{(n/2)-1} \cdot e^{-X/2}}{2^{n/2} \Gamma(n/2)} dx = 0.05 \quad \dots (9)$$

【0032】

【発明の効果】本発明によれば、観測条件を観測者が自由に決定できる場合に、従来に比べて信頼性が高く、しかも、簡単な構成で、短時間に、観測条件が所定の工程に影響を与えているか否かを判定できるデータ処理装置、その方法およびそのプログラムを提供することができる。

【図面の簡単な説明】

【図1】図1は、本発明の実施形態に係わるコンピュータの全体構成図である。

【図2】図2は、図1に示す²分布表データを説明するための図である。

【図3】図3は、図1に示すCPUの機能を説明するための図である。

【図4】図4は、図1に示すコンピュータが対象となる処理に用いられるデータを説明するための図である。

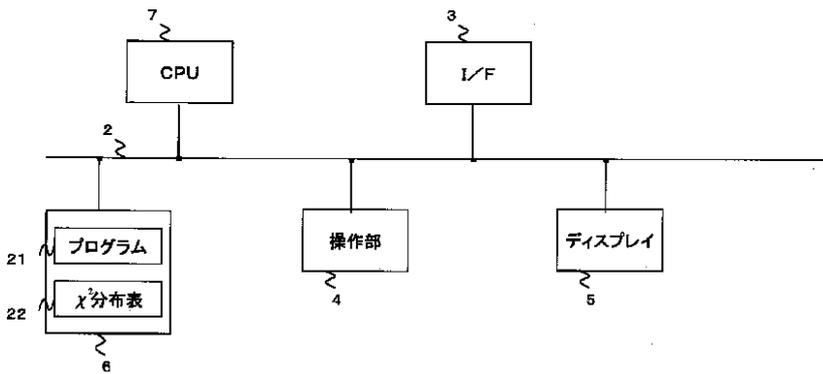
【図5】図5は、図1に示すコンピュータの動作例を説*

* 明するための図である。

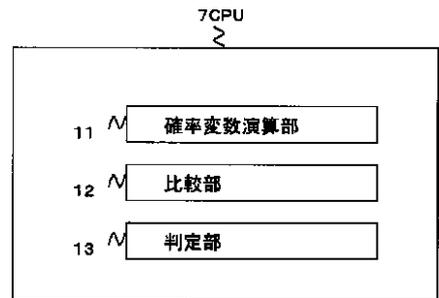
【符号の説明】

1...コンピュータ、2...バス、3...インタフェース、4...操作部、5...ディスプレイ、6...メモリ、7...CPU、11...確率変数演算部、12...比較部、13...判定部、21...プログラム、22...²分布表データ

【図1】



【図3】



1

【図2】

ϵ	.200	.100	.050	.025	.010	.005	.001
S-1							
1	1.64237	2.70554	3.84146	5.02389	6.63490	7.87944	10.8276
2	3.21888	4.60517	5.99146	7.37776	9.21034	10.5966	13.8155
3	4.64163	5.25139	7.81473	9.34840	11.3449	12.8382	16.2662
4	5.98862	7.77944	9.48773	11.1433	13.2767	14.6603	18.4668
5	7.28928	9.23636	11.0705	12.8325	15.0863	16.7496	20.5150

【図4】

(A)

条件\事象	E_1	E_2	E_3	E_4	合計
第1の条件	f_1	f_2	f_3	f_4	N
第2の条件	g_1	g_2	g_3	g_4	N

(B)

薬品有無\等級	1	2	3	4	合計
未使用	254	590	520	136	1500
使用	202	595	540	163	1500

【図5】

