



HOKKAIDO UNIVERSITY

Title	X線フィルム濃度分布の低濃度領域における統計的モデル
Author(s)	花田, 博之; 加藤, 浩; 関之山, 勝博 他
Citation	北海道大学医療技術短期大学部紀要, 3, 71-76
Issue Date	1990-10
Doc URL	https://hdl.handle.net/2115/37513
Type	departmental bulletin paper
File Information	3_71-76.pdf



X線フィルム濃度分布の低濃度領域に おける統計的モデル

花田 博之・加藤 浩・関之山勝博
下妻 光夫・山口 成厚

Statistical Model of X-ray Film Density Distribution after Irradiation of Weak X-rays

Hiroyuki Hanada, Hiroshi Kato, Katsuhiro Sekinoyama,
Mitsuo Shimosuma and Shigeatsu Yamaguchi

Abstract

A statistical model of the x-ray film density distribution of a screen-film system is described for the purpose of obtaining the information on the film density after the irradiation of weak x-rays. To determine the optimum statistical model, Akaike information criterion (AIC) is used.

In this study, the gamma distribution model was assumed as a result of the observation about the experimental data, and then the parameter of this model was determined by means of AIC. This optimum statistical model fitted exactly the density distribution.

要 旨

弱い X 線によるフィルム濃度分布の統計学的性質を調べた。低濃度領域のフィルム濃度分布は正規分布からのずれが大きいことが知られているが、適正な統計学的モデルを決定できれば、X 線センチメトリにたいして有用な情報を与える。本研究ではこれらモデルの決定に数理統計学における情報量基準の AIC を用いた。

実験で得られたデータを解析した結果、統計学的モデルとしてガンマ分布を仮定し、AIC によりそのパラメータを決定した。得られた最適モデルは実験データによく適合し、一般的なモ

デルの正規分布と比較しても適切なモデルであることがわかった。

1. はじめに

X 線撮影系において増感紙フィルムシステムのフィルム濃度分布の統計学的性質を知ることとは、X 線センチメトリに対して有用な情報を与える。特に低濃度領域における X 線フィルム濃度分布の統計学的性質が明らかになれば、スリット撮影による線像強度分布からレスポンス関数をもとめる方法に対して重要な役割をはたすことになる。統計学的性質の具体化は統計的モデルを数学的にうちたてることである。本

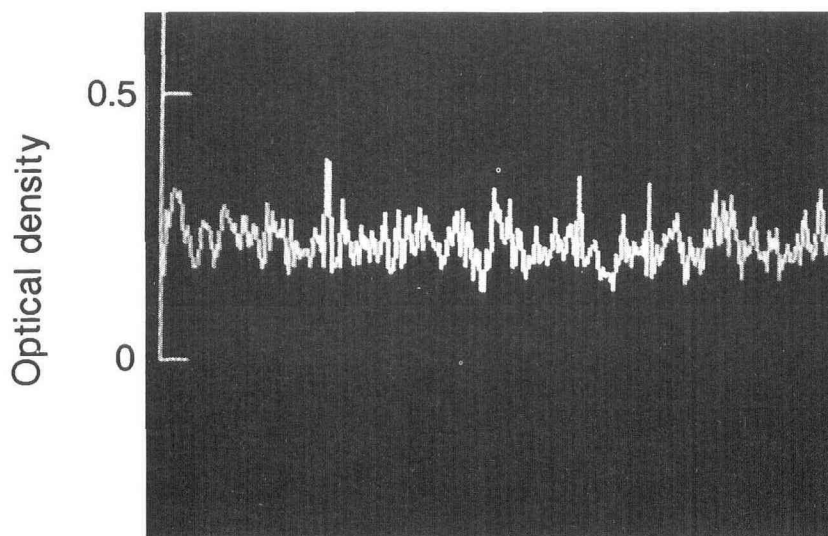


図1 濃度の波形 縦軸の目盛りは濃度値

研究では弱いX線を照射したときフィルム上に現れる低濃度領域の濃度分布に、正規分布とガンマ分布をあてはめ、数理統計学における情報量基準のAICを用いて最適モデルを決定した。

2. 濃度データの収集

濃度分布のフィルム試験片は以下の装置、材料、条件で作成した。X線装置、増感紙、X線フィルムには、それぞれ島津ED150L、KYOKKO BM-III、サクラAを使用した。照射条件は日本放射線技術学会総合委員会報告¹⁾の内容におおむね準拠させ、管電圧、管電流、照射時間をそれぞれ80 kV、30 mA、0.04 sとした。付加フィルタは3.0 mmAl+0.5 mmCu、X線管焦点・フィルム間距離は320 cmである。

濃度測定にはユニオン光学製マイクロフォトメータMPM-2型を用いた。測定範囲はフィルム中央部20×20 mmの正方形部分で、測定スリットの大きさは試料面で50×10 μmである。

濃度データの収集は、濃度に対応したマイクロフォトメータの出力電圧を、ADCを経てパソコンへ入力し、以後の計算処理に使用した。データ数は12300個、平均濃度は0.225、分散は

0.0023である。濃度が変動する様子を図1に示す。また濃度分布の様子をヒストグラムにしたものを図2に示す。縦軸は確率密度、横軸は濃度、0.225は濃度の平均値、各区分幅は0.117である。図2を観察すれば、一見して低濃度側(左側)に偏った左右非対称な分布になっていることがわかる。

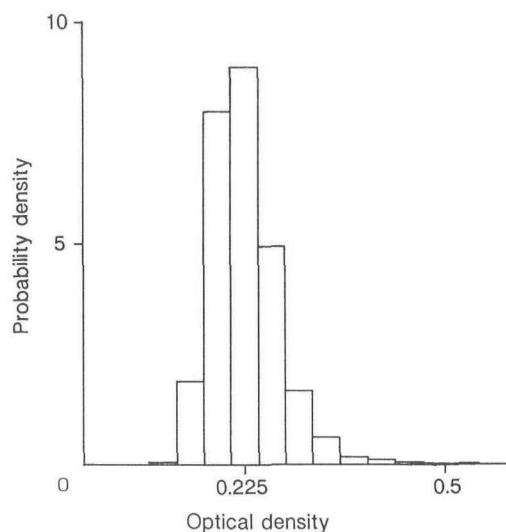


図2 濃度分布のヒストグラム
縦軸は確率密度、横軸は濃度

3. 統計的モデルと AIC

一般に実験で得られる統計的なデータ群を支配する真の分布則は未知である。通常それらデータ群に2つの統計的モデルをあてはめ、一つを帰無仮説、もう一つを対立仮説として有意水準5%の棄却域でモデルの適不適を判定することがひろく行なわれている。しかし仮説そのものに主観が入ることと、有意水準の設定値にあいまいさがつきまとうことが難点である。これに対して近年赤池²⁾³⁾により提唱された情報量基準は統計的モデルを評価する方法としてすぐれ、種々の分野で広く利用され成功をおさめている⁴⁾⁵⁾。この情報量基準は Akaike information criterion と呼ばれ、通常 AIC と記される。

AIC は以下のようにして定義される⁵⁾⁶⁾。

観測データ群を $x_1, \dots, x_n$, その統計的モデルの密度関数を $f(x|\theta)$ とする。 $\theta = \theta_1, \theta_2, \dots$ は密度関数を規定するパラメータである。例えば統計的モデルを正規分布と仮定すれば、平均値と分散が θ_1 と θ_2 に相当する。

確率変数 x が独立な場合

$$L(\theta) = f(x_1|\theta)f(x_2|\theta)\cdots f(x_n|\theta) \quad (1)$$

を尤度と呼び、その自然対数をとった

$$\ell(\theta) = \sum_{i=1}^n \log f(x_i|\theta) \quad (2)$$

を対数尤度と呼ぶ。ここで n は観測データ数である。 $\ell(\theta)$ は一連の観測値 $x_1, \dots, x_n$ で定まる θ の関数である。 θ に適当な値を与えれば $\ell(\theta)$ を最大にすることができ、そのときの $\theta = \hat{\theta}$ を最尤推定量、 $\ell(\hat{\theta})$ を最大対数尤度という。

AIC は

$$\begin{aligned} \text{AIC} = & -2 \times (\text{モデルの最大対数尤度}) \\ & + 2 \times (\text{モデルのパラメータ数}) \end{aligned} \quad (3)$$

で定義される⁶⁾⁷⁾⁸⁾。

いくつかのモデルがあるとき、それぞれについて AIC を計算して最小の AIC をもつモデルが最適なモデルと考える²⁾³⁾⁶⁾⁷⁾⁸⁾。

本研究では濃度データに適用する統計的モデルとして、最も一般的なモデルである正規分布と、図2のように左に偏った分布を表わすガンマ分布とを仮定して、それぞれについて AIC 最小のモデルを決定し、最後に両者を比較して最適モデルを決定した。

4. 最適モデルの計算方法

4.1 正規分布モデルの AIC

正規分布の密度関数は

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2} \quad (4)$$

のように表わされる。ここで x は濃度、 μ は濃度の平均値、 σ^2 は分散である。

データ $x_1, \dots, x_n$ に関する対数尤度は

$$\ell(\mu, \sigma^2) = \sum_{i=1}^n \log \left\{ \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x_i-\mu)^2/2\sigma^2} \right\} \quad (5)$$

である。 $\ell(\mu, \sigma^2)$ を μ と σ^2 で微分して0とおき、得られた連立方程式を解けば、 μ と σ^2 の最尤推定量が得られる。それらはそれぞれ

$$\hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i \quad (6)$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \left(\frac{1}{n} \sum_{i=1}^n x_i \right)^2 \quad (7)$$

である。これらを(5)式に代入すると最大対数尤度

$$\ell(\hat{\mu}, \hat{\sigma}^2) = -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \hat{\sigma}^2 - \frac{n}{2} \quad (8)$$

が得られる。このモデルのパラメータ数が μ と σ^2 の2つであることを考慮して(8)式を(3)式に代入すると、AIC は

$$\text{AIC} = n(\log 2\pi + \log \hat{\sigma}^2 + 1) + 4 \quad (9)$$

となる。

4.2 ガンマ分布モデルの AIC

ガンマ分布の密度関数は

$$g(x|\alpha, \beta) = \frac{1}{\beta^{\alpha+1} \Gamma(\alpha+1)} x^{\alpha} e^{-x/\beta} \quad (10)$$

のように表わされる。ここで x は濃度, $\alpha > 0$, $\beta > 0$ はパラメータ, $\Gamma(\alpha+1)$ はガンマ関数で

$$\Gamma(\alpha+1) = \int_0^{\infty} s^{\alpha} e^{-s} ds \quad (11)$$

で定義される。正規分布の場合と同じように対数尤度 $\sum_{i=1}^n \log g(x_i|\alpha, \beta)$ を α と β で微分して 0 とおくと連立方程式が得られるが, ガンマ関数の存在のため連立方程式の一方に α に関する積分方程式が現れ, 最尤推定量をもとめるのが困難になる。

そこで α を正の整数値をとる定数とみなすことにすれば, ガンマ関数は

$$\Gamma(\alpha+1) = \alpha! \quad (12)$$

となり, (10) 式の密度関数は

$$g(x|\beta) = \frac{1}{\beta^{\alpha+1} \alpha!} x^{\alpha} e^{-x/\beta} \quad (13)$$

となる。この場合, 密度関数を規定するパラメータは β のみである。これより対数尤度は

$$\ell(\beta) = \sum_{i=1}^n \left\{ \alpha \log x_i - \frac{x_i}{\beta} - (\alpha+1) \log \beta - \log \alpha! \right\} \quad (14)$$

となり, $\ell(\beta)$ を β で微分して 0 とおいて β の最尤推定量を求めると

$$\hat{\beta} = \frac{1}{n(\alpha+1)} \sum_{i=1}^n x_i \quad (15)$$

が得られる。(15) 式を (14) 式に代入して最大対数尤度を計算すると

$$\ell(\hat{\beta}) = \alpha \sum_{i=1}^n \log x_i - n(\alpha+1) - n(\alpha+1) \log \hat{\beta} - n \log \alpha! \quad (16)$$

となり, パラメータ数が 1 であることを考慮して (16) 式を (3) 式に代入して AIC を求めると

$$\begin{aligned} \text{AIC} = & -2\alpha \sum_{i=1}^n \log x_i + 2n(\alpha+1)(1 + \log \hat{\beta}) \\ & + 2n \log \alpha! + 2 \end{aligned} \quad (17)$$

が得られる。

本研究では $\alpha = 1, 2, \dots, 1000$ についてそれぞれ AIC を求め, それらの中の最小の AIC をもつモデルをガンマ分布の最適モデルとした。

5. 結 果

正規分布モデルに関する (6), (7), (8), (9) 式において n にデータ数 12300 を与えて, 平均値と分散の最尤推定量ならびに AIC を計算すると以下ようになる。

$$\begin{aligned} \hat{\mu} &= 0.225 \\ \hat{\sigma}^2 &= 0.0023 \\ \text{AIC} &= -39889.0 \end{aligned} \quad (18)$$

ガンマ分布モデルの AIC は, $\alpha = 1$ から $\alpha = 1000$ まで 1 ずつ増加させて計算した。AIC はマイナスの値からプラスの値まで大きく変化するので, 全体の増減の様子をわかりやすくするため, 各 AIC に 42000 を加えてその常用対数をとったものを図 3 に示す。これより $\alpha = 10$ から

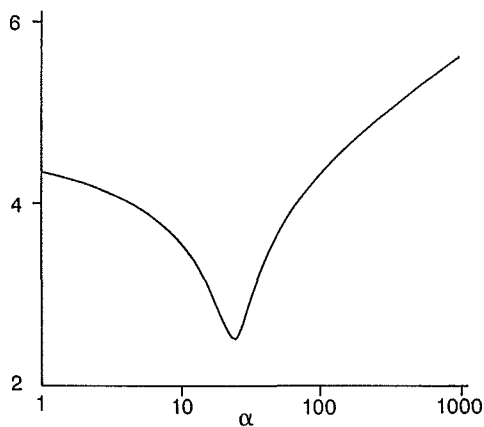


図3 α の変化に対する AIC の増減

$\alpha = 100$ までの区間内に最小の AIC が存在することがわかる。表 1 は $\alpha = 11$ から $\alpha = 40$ までの区間の AIC である。図 3 と表 1 より $\alpha = 24$ のときの AIC が最小である。この場合の β の最尤推定量と AIC を以下に示す。

表1 ガンマ分布の α とAICの値

α	AIC	α	AIC
11	-39043.3	26	-41639.4
12	-39542.7	27	-41591.0
13	-39967.3	28	-41526.9
14	-40327.4	29	-41447.8
15	-40631.6	30	-41355.5
16	-40886.6	31	-41249.9
17	-41098.0	32	-41131.7
18	-41271.0	33	-41002.7
19	-41409.1	34	-40862.7
20	-41515.7	35	-40712.5
21	-41594.0	36	-40552.9
22	-41646.6	37	-40384.4
23	-41675.6	38	-40206.7
24	-41683.1	39	-40021.5
25	-41670.2	40	-39828.5

$$\hat{\beta} = 0.00898$$

$$AIC = -41683.1 \quad (19)$$

以上により(4)式と(13)式に(18)式と(19)式でもとめた最尤推定量を代入すると、正規分布モデルとガンマ分布モデルの密度関数が決定される。それらを $f(x)$ 、 $g(x)$ とすれば

$$f(x) = \frac{1}{\sqrt{0.0046\pi}} e^{-(x-0.225)^2/0.0046} \quad (20)$$

$$g(x) = \frac{1}{0.00898^{25} \cdot 24!} x^{24} e^{-x/0.00898} \quad (21)$$

である。

(20)式と(21)式の分布を図2のヒストグラムに重ねてかいたのが図4と図5である。

6. 考 察

本研究の対象とした濃度は、フィルム特性曲線上の足部に相当する領域にあり、フィルムベースとカブリによるベース濃度にきわめて近い値を持つ。これはまた線像強度分布のすそに相当する領域の濃度でもある。これら領域の濃度分布は横田らの報告⁹⁾¹⁰⁾¹¹⁾で正規分布からずれてくることが知られている。図2によれば本実験の場合も分布の重みが低濃度側で大きくなっている。ところで観測値が正の値のみをと

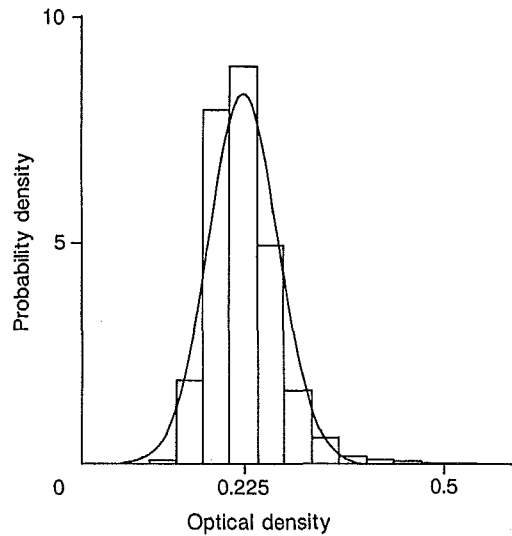


図4 濃度分布と正規分布の密度関数
縦軸は確率密度、横軸は濃度

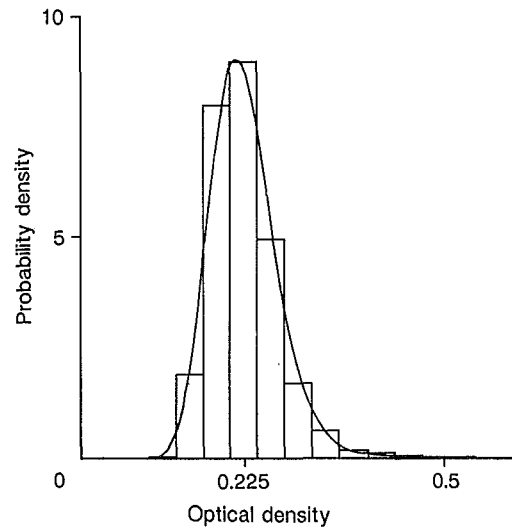


図5 濃度分布とガンマ分布($\alpha=24$)の密度関数
縦軸は確率密度、横軸は濃度

る分布で数学的にも扱いやすいものにガンマ分布がある。濃度値は常に正であり、本研究でガンマ分布を統計的モデルに選んだ理由はここにある。

統計的モデルの選択の基準となるAICは、その絶対値に意味があるのではなく、他のモデル

と比較したときの相対値のみに意味があり、値の小さいモデルほど良いモデルだと判定される²⁾³⁾⁶⁾⁷⁾⁸⁾。

表1によればガンマ分布モデルにおいて、 $\alpha=13$ から $\alpha=39$ までの範囲のすべての AIC は、正規分布モデルの AIC の -39889 より小さい。このことと図3の極小値(最小値でもある)がただひとつ存在することと合わせて考えれば、本実験でのガンマ分布モデルの適用の妥当性が理解できる。さらに図4と図5を比較すると正規分布モデルに対するガンマ分布モデル($\alpha=24$)の優位性は明かである。

ところで4.2で記したように積分方程式の取り扱いを回避するため、 α を正の連続量をとるパラメータとして扱うかわりに、正の整数に限定した定数とみなして処理した。このことの妥当性は、 α の増加とともに AIC が単調減少し、最小値に達したのち単調増加に転ずる様子を表わした図3により明かである。特定の α に対して特異な現象は起こってはず、最小値付近の α の増減に対する AIC の増減もわずかであるからである。

7. ま と め

以下の結果を得た。

1) 弱い X 線によるフィルム濃度分布についてのガンマ分布モデルの決定法を導いた。

2) 本実験の条件では、X 線フィルム濃度分布の統計的モデルとして、 $\alpha=24$, $\beta=0.00898$ のガンマ分布が最適であった。

最後に、本論文の図表作成にあたって有益な助言を下された本学医療技術短期大学部診療放射線技術学科荒井博史氏に感謝いたします。

文 献

- 1) 日本放射線技術学会総合委員会：増感紙フィルムシステムの X 線センシトメトリについて，日放技学誌，40(1)，59-78，1984
- 2) H. Akaike : Information theory and an extension of the maximum likelihood principle, 2nd Inter. Symp. on Information Theory, 267-281, 1973
- 3) H. Akaike : A new look at the statistical model identification, IEEE Trans. Autom. Contr., 716-723, 1974
- 4) 赤池弘次，中川東一郎：ダイナミックシステムの統計的解析と制御，94-122，1972，サイエンス社
- 5) 赤池弘次，他：モデルによってデータを測る，数理科学，No. 213，7-66，1981
- 6) 坂元慶行，石黒真木夫，北川源四郎：情報量統計学，27-64，1983，共立出版
- 7) 赤池弘次：情報量基準 AIC とは何か，数理科学，No. 153，5-11，1976
- 8) 赤池弘次：統計的検定の新しい考え方，数理科学，No. 198，51-57，1979
- 9) 横田豊，他：X 線写真濃度分布の検討（第1報），日放技学誌，44(8)，860，1988
- 10) 横田豊，他：X 線写真濃度分布の検討（第2報），日放技学誌，45(2)，183，1989
- 11) 横田豊，他：X 線写真濃度分布の検討（第3報），日放技学誌，45(8)，1044，1989