

Statistical prediction regions and limits

筑波大・数理解析 日吉 貴之 (Takayuki Hiyoshi)
Graduate School of Pure and Applied Science
University of Tsukuba
筑波大・数理解析 赤平 昌文 (Masafumi Akahira)
Graduate School of Pure and Applied Science
University of Tsukuba

1 はじめに

数理統計学において未観測の確率変数の値を予測することは基本的な問題である。まず、観測可能な X の値に基づく未観測の確率変数 Y に関する何らかの判断を下すルールを定めることを考える ([T75], [G93])。すなわち、可能な結論の集合を Π とするとき、 X の実現値 x に対して Π の要素 $p = \pi(x)$ を定めるルール π を求めることである。また Y の値が y であるとき、下した結論 $p \in \Pi$ の損失を表す損失関数 $L(y, p)$ が定められているとする。そのとき、予測ルール π に対応するリスク $R(\theta, \pi) := E_\theta[L(Y, \pi(x))]$ をなるべく小さくするように、 π を定めることが目的となる。これはベイズ的観点から、データが与えられたとき未観測の確率変数の事後予測密度による事後リスクを最小にする π を求めればよいが、 θ の事前分布が存在するか、あるいは何らかの理由によってそれが想定される場合に限られる。 θ について特定の事前分布を想定する十分な根拠がない場合は不偏性や不変性などの制限をつけたり、また最近では予測尤度の概念によって予測のための非ベイズ的尤度の観点からの試みがなされている。

本論では、まず予測尤度の構成と概念について Bjørnstad[B90] に従って考察し、その具体例を挙げる。次に、上側予測限界 $u(X)$ について考え、 X を与えたときの Y の $u(X)$ による下側確率の分散に対する上界を導出するとともに、ノンパラメトリックの場合、すなわち分布型が未知の場合や分布型が既知の場合にそれらの分散を比較し、どれだけ精度が良くなっているかについて考察する。

2 統計的予測尤度

本節において、予測尤度の定義について述べて、具体的な例を挙げる。そして予測尤度の性質およびそれに基づく予測域について論じる。

2.1 設定

観測可能な確率ベクトル $\mathbf{X} = (X_1, \dots, X_n)$ 、未観測確率ベクトル $\mathbf{Y} = (Y_1, \dots, Y_m)$ 、 Z を $\sum_{i=1}^m Y_i$ や \mathbf{Y} 自身などの \mathbf{Y} の関数を予測すべき量とする。また、 (\mathbf{X}, \mathbf{Y}) は同時確率密度関数 (j.p.d.f.) $f_{\mathbf{X}, \mathbf{Z}}^\theta(\mathbf{x}, z)$ を持つとする。ここで θ は未知の母数ベクトルとする。ここで、観測可能な確率ベクトル \mathbf{X} に基づいて、未観測な確率変数 Z を予測する問題を考える。この予測問題においては、 Z, θ は未知で、 θ を局外母数と見なして Z に関する情報をいかに得るかが本質的である。そこで、 (Z, θ) の尤度関数 $l_{\mathbf{x}}(z, \theta) := f_{\mathbf{X}, \mathbf{Z}}^\theta(\mathbf{x}, z)$ から何らかの方法で θ を消去した Z の尤度関数 $L(z|\mathbf{x})$ を予測尤度関数と呼ぶ。

本論では予測尤度として周辺最大予測尤度

$$L_p(z|\mathbf{x}) = \sup_{\theta} f_{\mathbf{X}, \mathbf{Z}}^\theta(\mathbf{x}, z) = l_{\mathbf{x}}(z, \hat{\theta})$$

と一般一様事前密度に関するベイズ的事後予測密度

$$f_0(z|\mathbf{x}) = k(\mathbf{x}) \int_{-\infty}^{\infty} l_{\mathbf{x}}(z, \theta) d\theta$$

を取り上げ、その加法性等について論じる。ただし、 $\hat{\theta}$ は θ の最尤推定量 (MLE)、 $k(\mathbf{x})$ は規格化定数とする。

2.2 既知の分散をもつ正規分布の場合

まず、 $X_1, \dots, X_n, Y_1, \dots, Y_m$ をたがいに独立にいずれも正規分布 $N(\theta, \sigma_0^2)$ に従う確率変数とする。ただし σ_0^2 は既知とする。このとき $Z := \sum_{i=1}^m Y_i/m$ とすると、 $L_p(z|\mathbf{x})$ と $f_0(z|\mathbf{x})$ は正規分布

$$N\left(\bar{x}, \left(\frac{1}{m} + \frac{1}{n}\right) \sigma_0^2\right)$$

に従う確率変数の p.d.f. L_0 と同様になる。ただし、 $\mathbf{x} := (x_1, \dots, x_n)$, $\bar{x} = (1/n) \sum_{i=1}^n x_i$ とする。実際 $\mathbf{X} := (X_1, \dots, X_n)$ の同時 (j.)p.d.f. は \mathbf{X} と Z が独立であるから

$$f_{\mathbf{X}}^{\theta}(\mathbf{x}) = (2\pi)^{-\frac{n}{2}} \sigma_0^{-n} \exp\left\{-\frac{1}{2\sigma_0^2} \sum_{i=1}^n (x_i - \theta)^2\right\}$$

になる。ただし、 $\theta = (\theta, \dots, \theta)$ とする。また Z の p.d.f. は

$$f_Z^{\theta}(z) = \frac{\sqrt{m}}{\sqrt{2\pi}\sigma_0} e^{-\frac{m(z-\theta)^2}{2\sigma_0^2}}$$

となるので、 (\mathbf{X}, Z) の j.p.d.f. は

$$\begin{aligned} f_{\mathbf{X}, Z}^{\theta}(\mathbf{x}, z) &= f_{\mathbf{X}}^{\theta}(\mathbf{x}) \cdot f_Z^{\theta}(z) \\ &= \sqrt{m} \left(\frac{1}{\sqrt{2\pi}\sigma_0}\right)^{n+1} \exp\left[-\frac{1}{2\sigma_0^2} \left\{\sum_{i=1}^n (x_i - \theta)^2 + m(z - \theta)^2\right\}\right] \end{aligned}$$

になる。このとき

$$\begin{aligned} \sum_{i=1}^n (x_i - \theta)^2 + m(z - \theta)^2 &= (m+n)\theta^2 - 2(n\bar{x} + mz)\theta + \sum_{i=1}^n x_i^2 + mz^2 \\ &= (m+n) \left\{\theta - \frac{n\bar{x} + mz}{m+n}\right\}^2 - \frac{(n\bar{x} + mz)^2}{m+n} + \sum_{i=1}^n x_i^2 + mz^2 \end{aligned}$$

であるから、 (\mathbf{X}, Z) に基づく θ の MLE は

$$\hat{\theta} = \frac{n\bar{x} + mz}{m+n}$$

となるので

$$\begin{aligned} L_p(z|\mathbf{x}) &= \left(\frac{1}{\sqrt{2\pi}\sigma_0}\right)^n \frac{\sqrt{m+n}}{\sqrt{n}} \exp\left[-\frac{1}{2\sigma_0^2} \left\{-\frac{n^2\bar{x}^2}{m+n} + \sum_{i=1}^n x_i^2 - \left(\frac{mn}{m+n}\bar{x}\right)^2\right\}\right] \\ &\quad \cdot \frac{\sqrt{mn}}{\sqrt{2\pi}\sigma_0\sqrt{m+n}} \exp\left\{-\frac{1}{2\sigma_0^2} \frac{mn}{m+n} (z - \bar{x})^2\right\} \end{aligned}$$

$\propto L_0$

となる。ただし、 α は比例関係を表わす。一方

$$\begin{aligned} f_0(z|\mathbf{x}) &\propto \int l_{\mathbf{x}}(z, \theta) d\theta \\ &= \sqrt{m} \left(\frac{1}{\sqrt{2\pi}\sigma_0} \right)^{n+1} \exp \left[-\frac{1}{2\sigma_0^2} \left\{ mz^2 - \frac{(n\bar{x} + mz)^2}{m+n} + \sum_{i=1}^n x_i^2 \right\} \right] \cdot L_0 \end{aligned}$$

となり、規準化すると $f_0 = L_0$ となる。したがって $L_p(\cdot|\mathbf{x})$ と $f_0(\cdot|\mathbf{x})$ は本質的に同じ予測尤度になる。

2.3 ベルヌーイ分布の場合

まず、 $X_1, \dots, X_n, Y_1, \dots, Y_m$ をたがいに独立にいずれもベルヌーイ分布 $\text{Ber}(\theta)$ に従う確率変数とする。このとき、 $Z := \sum_{i=1}^m Y_i, S := \sum_{j=1}^n X_j$ とすると、周辺最大予測尤度 $L_p(z|\mathbf{x})$ は

$$\binom{m}{z} (s+z)^{s+z} (m+n-s-z)^{m+n-s-z}$$

に比例し、一般一様事前分布に関するベイズ的事後予測密度は

$$f_0(z|\mathbf{x}) = \left\{ \frac{\binom{m}{z} \binom{n}{s}}{\binom{m+n}{s+z}} \right\} \cdot \frac{n+1}{m+n+1} \quad (z = 0, 1, \dots, m)$$

になる。すなわち、 Z の事後分布は負の超幾何分布になる。実際、 $\mathbf{X} := (X_1, \dots, X_n)$ の同時確率量関数 (j.p.m.f.) は

$$f_{\mathbf{X}}^{\theta}(\mathbf{x}) = \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n-\sum_{i=1}^n x_i} = \theta^s (1-\theta)^{n-s} \quad (s = 0, 1, \dots, n)$$

となり、 Z の p.m.f. は

$$f_Z^{\theta}(z) = \binom{m}{z} \theta^z (1-\theta)^{m-z} \quad (z = 0, 1, \dots, m)$$

となるので、 (\mathbf{X}, Z) の j.p.m.f. は \mathbf{X} と Z が独立であるから

$$f_{\mathbf{X}, Z}^{\theta}(\mathbf{x}, z) = f_{\mathbf{X}}^{\theta}(\mathbf{x}) \cdot f_Z^{\theta}(z) = \theta^s (1-\theta)^{n-s} \binom{m}{z} \theta^z (1-\theta)^{m-z} = \binom{m}{z} \theta^{s+z} (1-\theta)^{m+n-s-z}$$

となる。このとき

$$\log f_{\mathbf{X}, Z}^{\theta}(\mathbf{x}, z) = \log \binom{m}{z} + (s+z) \log \theta + (m+n-s-z) \log(1-\theta)$$

となるから

$$\frac{\partial}{\partial \theta} \log f_{\mathbf{X}, Z}^{\theta}(\mathbf{x}, z) = \frac{s+z}{\theta} - \frac{m+n-s-z}{1-\theta} = 0$$

より、 (\mathbf{X}, Z) に基づく θ の MLE は

$$\hat{\theta} = \frac{s+z}{m+n}$$

となる。よって、周辺最大予測尤度は

$$L_p(z|\mathbf{x}) = \sup_{\theta} f_{\mathbf{X},Z}^{\theta}(\mathbf{x}, z) \propto \binom{m}{z} (s+z)^{s+z} (m+n-s-z)^{m+n-s-z}$$

となる。一方、一般一様事前密度に関するベイズ的事後予測密度は

$$f_0(z|\mathbf{x}) \propto \int_0^1 f_{\mathbf{X},Z}^{\theta}(\mathbf{x}, z) d\theta = \left\{ \binom{m}{z} / \binom{m+n}{s+z} \right\} \cdot \frac{1}{m+n+1}$$

であるから

$$f_0(z|\mathbf{x}) = k(\mathbf{x}) \left\{ \binom{m}{z} / \binom{m+n}{s+z} \right\} \cdot \frac{1}{m+n+1}$$

と書ける。ただし、 $k(\mathbf{x})$ は規準化定数とする。ここで、規準化定数 $k(\mathbf{x})$ を求めると、 $k(\mathbf{x}) = \binom{n}{s} \cdot (n+1)$ となり

$$f_0(z|\mathbf{x}) = \left\{ \binom{m}{z} \binom{n}{s} / \binom{m+n}{s+z} \right\} \cdot \frac{n+1}{m+n+1}$$

になる。よって、 Z の事後分布は負の超幾何分布になることがわかる。

2.4 予測域

次に、周辺最大予測尤度、ベイズ的最大予測尤度によって予測域を、通常未知母数の信頼域と同様に考えることができる ([A03], [H00], [HH98])。予測尤度 $L(z|\mathbf{x})$ に対して

$$R_{\mathbf{x}}(c) := \{z : L(z|\mathbf{x}) \geq c\}$$

とおいて、 $0 < \alpha < 1$ について

$$\int_{R_{\mathbf{x}}(c)} L(z|\mathbf{x}) dz = 1 - \alpha \quad (\text{連続型の場合}), \quad \sum_{z \in R_{\mathbf{x}}(c)} L(z|\mathbf{x}) = 1 - \alpha \quad (\text{離散型の場合})$$

となる $R_{\mathbf{x}}(c)$ を信頼係数 $1 - \alpha$ の予測尤度 $L(z|\mathbf{x})$ による最大予測域という。 $L(z|\mathbf{x})$ が z の関数として単峰形であれば、 $R_{\mathbf{x}}(c)$ は区間になり、予測区間を得る。周辺最大予測尤度やベイズ的事後予測尤度についても、 $L(z|\mathbf{x})$ の代わりに $L_p(z|\mathbf{x})$, $f_0(z|\mathbf{x})$ とすれば同様にできる。

最初に第 2.2 節で扱った正規分布において、 $n = 25$, $m = 5$ の場合を考える。この場合、規準化した周辺最大予測尤度と一般一様事前密度に関するベイズ的事後予測尤度は、ともに $N(\bar{x}, ((m+n)/(mn))\sigma_0^2)$ の p.d.f.

$$\frac{\sqrt{mn}}{\sqrt{2\pi\sigma_0\sqrt{m+n}}} \exp \left\{ -\frac{1}{2\sigma_0^2} \frac{mn}{m+n} (z - \bar{x})^2 \right\}$$

となるので、これを用いて $Z = (1/5) \sum_{j=1}^5 Y_j$ の最大予測域を求める。

まず真値 θ_0 を与えた $N(\theta_0, \sigma_0^2)$ から大きさ 25 の無作為標本を抽出し、 Z の最大予測域を求める。このとき、この操作を 1000 回反復し、予測限界の値を平均して最大予測区間を作る。実際 $\theta_0 = 0, \sigma_0^2 = 1$ の最大予測区間は表 2.1, $\theta_0 = 1, \sigma_0^2 = 4$ の最大予測区間は表 2.2 となる。

$1 - \alpha$	予測区間
0.95	(-0.9545, 0.9658)
0.90	(-0.8002, 0.8115)
0.80	(-0.6222, 0.6335)
0.50	(-0.3248, 0.3361)

表 2.1

$1 - \alpha$	予測区間
0.95	(-0.9286, 2.9122)
0.90	(-0.6198, 2.6034)
0.80	(-0.2638, 2.2475)
0.50	(0.3309, 1.6527)

表 2.2

次に, $X_1, \dots, X_n, Y_1, \dots, Y_m$ をたがいに独立にいずれも指数分布 $\text{Exp}(\lambda)$ に従う確率変数とする場合を考える. ここで, $Z := (1/m) \sum_{j=1}^m Y_j$ とするとき, その周辺最大予測尤度は

$$\frac{m^m}{\Gamma(m)} e^{-(n+m)} \left(\frac{n+m}{n\bar{x} + mz} \right)^{n+m} z^{m-1}$$

となる. これを規準化して用いて, $n = 25, m = 5$ の場合の $Z = (1/5) \sum_{j=1}^5 Y_j$ の最大予測域を考える.

まず真値 λ_0 を与えた $\text{Exp}(\lambda_0)$ から大きさ 25 の無作為標本を抽出し, Z の最大予測域を求める. このとき, この操作を 1000 回反復し, 予測限界の値を平均して最大予測区間を作る. 実際 $\lambda_0 = 1$ の最大予測区間は表 2.3, $\lambda_0 = 2$ の最大予測区間は表 2.4 となる.

$1 - \alpha$	予測区間
0.95	(0.3012, 2.6696)
0.90	(0.3491, 2.1715)
0.80	(0.4543, 1.7900)
0.50	(0.6997, 1.5478)

表 2.3

$1 - \alpha$	予測区間
0.95	(0.6009, 5.3267)
0.90	(0.7005, 4.3574)
0.80	(0.8279, 3.4350)
0.50	(1.1166, 2.3950)

表 2.4

最後に, 第 2.3 節で扱ったベルヌーイ分布において, $n = 25, m = 5$ の場合を考える. この場合, 一般一様事前密度に関するベイズ的事後予測密度は負の超幾何分布の p.d.f.

$$\left\{ \binom{m}{z} \binom{n}{s} / \binom{m+n}{s+z} \right\} \cdot \frac{n+1}{m+n+1} \quad (z = 0, 1, \dots, m)$$

となるので, これを用いて $Z = \sum_{j=1}^5 Y_j$ の最大予測域を考える.

ここで, Z の事後分布は負の超幾何分布で, これは離散型になるので, 非負の整数 c_1, c_2 ($c_1 < c_2$) について予測域を区間 $[c_1, c_2]$ の形で表せるが, 一般には予め与えられた信頼係数を達成できない. そこで, 信頼係数 $1 - \alpha$ の最大予測域 (区間) のランダム検定関数

$$\phi(Z) = \begin{cases} 1 & (c_1 \leq Z \leq c_2), \\ \gamma & (Z = c_1 - 1, c_2 + 1), \\ 0 & (\text{その他}) \end{cases}$$

を考える. ただし, 非負の整数 c_1, c_2 と $0 \leq \gamma \leq 1$ は $E(\phi) = 1 - \alpha$ となるように定める. しかし, このランダム検定は必ずしも実用的ではないので, 実際には次のように行う.

まず、確率変数 U を一様分布 $U(0,1)$ に従い、 Z とは独立とする。そして、

$$T := Z + U$$

とおくと

$$\begin{aligned} f_T(t) &= P\{T = t\} = P\{Z + U = t\} \\ &= P\{Z = z, U = t - z\} \\ &= P\{Z = z\}P\{U = t - z\} \\ &= f_Z(z)f_U(t - z) \\ &= f_Z(z) \quad (z \leq t < z + 1) \end{aligned}$$

となるので、予測域 $\{c_1 - \gamma \leq T \leq c_2 + 1 + \gamma\}$ をもつ非ランダム検定はランダム検定 ϕ と同等になる。

また、 $P\{Z = c_1 - 1\}$ と $P\{Z = c_2 + 1\}$ の比率を生かすために、 $P\{Z = c_1 - 1\}/P\{Z = c_2 + 1\} = \gamma_1/\gamma_2$ として、信頼係数 $1 - \alpha$ の最大予測域 (区間) のランダム検定関数

$$\phi^*(Z) = \begin{cases} 1 & (c_1 \leq Z \leq c_2), \\ \gamma_1 & (Z = c_1 - 1), \\ \gamma_2 & (Z = c_2 + 1), \\ 0 & (\text{その他}) \end{cases}$$

を考える。ただし、非負の整数 c_1, c_2 と $0 \leq \gamma_1, \gamma_2 \leq 1$ は $E(\phi^*) = 1 - \alpha$ となるように定める。このとき $T^* := Z + U$ として同様に考えると予測域は区間 $\{c_1 - \gamma_1 \leq T^* \leq c_2 + 1 + \gamma_2\}$ となる。

まず、母数 θ が $1/2$ の場合のベルヌーイ分布を考える。ここで、 $\text{Ber}(1/2)$ から 25 個の無作為標本を抽出し、 $S = \sum_{i=1}^{25} X_i = 12$ であったとする。 $Z = \sum_{j=1}^5 Y_j$ であるから、このときランダム検定関数として、表 2.5 のような $1 - \alpha, c_1, c_2, \gamma, \gamma_1, \gamma_2$ をとる ϕ と ϕ^* を考えれば、 T, T^* の予測区間は表 2.6 のようになる。

$1 - \alpha$	c_1	c_2	γ	γ_1	γ_2
0.95	1	4	0.4243	0.4802	0.3468
0.90	1	3	0.6827	0.2756	0.8196
0.80	1	3	0.1837	0.0741	0.2205
0.50	2	2	0.4319	0.3247	0.5013

表 2.5

$1 - \alpha$	T の予測区間	T^* の予測区間
0.95	(0.5757, 5.4243)	(0.5198, 5.3468)
0.90	(0.3173, 4.6827)	(0.7244, 4.8196)
0.80	(0.8163, 4.1837)	(0.9259, 4.2205)
0.50	(1.5681, 3.4319)	(1.6753, 3.5013)

表 2.6

同様にして、母数 θ が $1/3$ の場合のベルヌーイ分布を考える。ここで、 $\text{Ber}(1/3)$ から大きさ 25 の無作為標本を抽出し、 $S = \sum_{i=1}^{25} X_i = 10$ であったとする。 $Z = \sum_{j=1}^5 Y_j$ であるから、このときランダム検定関数として、表 2.7 のような $1 - \alpha, c_1, c_2, \gamma, \gamma_1, \gamma_2$ をとる ϕ と ϕ^* を考えると、 T, T^* の予測域は表 2.8 のようになる。

$1 - \alpha$	c_1	c_2	γ	γ_1	γ_2
0.95	0	3	0.4658	*	*
0.90	1	3	0.5389	0.6344	0.3789
0.80	1	3	0.0161	0.0190	0.0113
0.50	2	2	0.3856	0.4407	0.3062

表 2.7

$1 - \alpha$	T の予測区間	T^* の予測区間
0.95	(0 , 4.4658)	*
0.90	(0.4611 , 4.5389)	(0.3656 , 4.3789)
0.80	(0.9839 , 4.0161)	(0.9810 , 4.0113)
0.50	(1.6144 , 3.3856)	(1.5593 , 3.3062)

表 2.8

3 統計的予測限界

未観測の確率変数の予測限界を累積分布関数 (c.d.f.) を通して分散による評価を行い、パラメトリックモデルとノンパラメトリックモデルの場合の差異について考察する。

3.1 予測限界の分布関数の分散の上界

観測可能な確率変数を X_1, \dots, X_n , 未観測の確率変数を Y とし, それらがたがいに独立にいずれも (σ -有限測度 μ に関する) p.d.f. $f(x, \theta)$, c.d.f. $F(x, \theta)$ をもつ分布に従うとする。ただし, $\theta \in \Theta \subset \mathbf{R}^1$ で θ は未知とする。

いま, $0 < \alpha < 1$ とするとき, $\mathbf{X} := (X_1, \dots, X_n)$ に対して区間 $[u(\mathbf{X}), \infty)$ を定めて, 任意の $\theta \in \Theta$ について

$$P_{\mathbf{X}, Y}^{\theta} \{Y \leq u(\mathbf{X})\} = 1 - \alpha \quad (3.1)$$

とする。このとき, $u(\mathbf{X})$ を信頼係数 $1 - \alpha$ の上側予測限界という。

このとき, $1 - \alpha = F(\xi_{\theta}, \theta)$ となる ξ_{θ} はこの分布の上側 $100\alpha\%$ 点となる。ここで ξ_{θ} を $\hat{\xi}_{\theta}(\mathbf{X})$ で推定し, これを $u(\mathbf{X})$ とみなす。このとき, \mathbf{X} を与えたときの Y の $\hat{\xi}_{\theta}(\mathbf{X})$ による下側確率の分散に対する上界を求める。

まず, 任意の θ に対して

$$E_{\theta}[F(\hat{\xi}_{\theta}, \theta)] = 1 - \alpha + o(1)$$

になる。ただし, χ は定義関数とする。このとき, $\hat{\theta} = \hat{\theta}(\mathbf{X})$ を θ の一致推定量として, ξ_{θ} の推定量を $\hat{\xi}_{\theta}(\mathbf{X}) := \xi_{\hat{\theta}(\mathbf{X})}$ とする。また, $\sqrt{n}(\hat{\theta} - \theta)$ は漸近分布をもつとする。そこで, n が十分大きいとき

$$\begin{aligned} F(\hat{\xi}_{\theta}, \theta) &= F(\xi_{\theta}, \theta) + f(\xi_{\theta}, \theta)(\hat{\xi}_{\theta} - \xi_{\theta}) + o_p\left(\frac{1}{\sqrt{n}}\right) \\ &= 1 - \alpha + f(\xi_{\theta}, \theta)(\hat{\xi}_{\theta} - \xi_{\theta}) + o_p\left(\frac{1}{\sqrt{n}}\right) \end{aligned} \quad (3.2)$$

になる。一方, $F(\xi_{\theta}, \theta) = 1 - \alpha$ の両辺を θ で微分すると

$$\left(\frac{\partial \xi_{\theta}}{\partial \theta}\right) f(\xi_{\theta}, \theta) + \left[\frac{\partial}{\partial \eta} F(\xi_{\theta}, \eta)\right]_{\eta=\theta} = 0 \quad (3.3)$$

となる。また

$$\hat{\xi}_{\theta} - \xi_{\theta} = \xi_{\hat{\theta}} - \xi_{\theta} = \frac{\partial \xi_{\theta}}{\partial \theta}(\hat{\theta} - \theta) + o_p\left(\frac{1}{\sqrt{n}}\right) \quad (3.4)$$

となる。ここで, (3.1.3) の辺々に $(\hat{\theta} - \theta)$ をかけると

$$\left(\frac{\partial \xi_{\theta}}{\partial \theta}\right) (\hat{\theta} - \theta) f(\xi_{\theta}, \theta) + (\hat{\theta} - \theta) \left[\frac{\partial}{\partial \eta} F(\xi_{\theta}, \eta)\right]_{\eta=\theta} = 0$$

となり, (3.1.4) から

$$(\hat{\xi}_\theta - \xi_\theta)f(\xi_\theta, \theta) + (\hat{\theta} - \theta) \left[\frac{\partial}{\partial \eta} F(\xi_\theta, \eta) \right]_{\eta=\theta} = o_p \left(\frac{1}{\sqrt{n}} \right)$$

となり, したがって

$$f(\xi_\theta, \theta)(\hat{\xi}_\theta - \xi_\theta) = - \left[\frac{\partial}{\partial \eta} F(\xi_\theta, \eta) \right]_{\eta=\theta} (\hat{\theta} - \theta) + o_p \left(\frac{1}{\sqrt{n}} \right) \quad (3.5)$$

となる. さらに (3.1.5) を (3.1.2) に代入すると

$$F(\hat{\xi}_\theta, \theta) = 1 - \alpha - \left[\frac{\partial}{\partial \eta} F(\xi_\theta, \eta) \right]_{\eta=\theta} (\hat{\theta} - \theta) + o_p \left(\frac{1}{\sqrt{n}} \right)$$

になる. ここでまた, n が十分大きいとき, $V_\theta(\hat{\theta}) = 1/(nI(\theta)) + o(1/n)$ を満たすとする, と

$$\begin{aligned} V_\theta(F(\hat{\xi}_\theta, \theta)) &= \left\{ \left[\frac{\partial}{\partial \eta} F(\xi_\theta, \eta) \right]_{\eta=\theta} \right\}^2 V_\theta(\hat{\theta}) + o \left(\frac{1}{n} \right) \\ &= \frac{1}{nI(\theta)} \left\{ \left[\frac{\partial}{\partial \eta} F(\xi_\theta, \eta) \right]_{\eta=\theta} \right\}^2 + o \left(\frac{1}{n} \right) \end{aligned} \quad (3.6)$$

になる. ただし, $I(\theta)$ は X のもつ θ に関する Fisher 情報量とする. さらに, 微分と積分が順序交換可能であると仮定すると

$$\begin{aligned} \frac{\partial}{\partial \eta} F(\xi_\theta, \eta) &= \frac{\partial}{\partial \eta} \int_{-\infty}^{\xi_\theta} f(x, \eta) d\mu(x) \\ &= \int_{-\infty}^{\xi_\theta} \frac{\partial}{\partial \eta} f(x, \eta) d\mu(x) \end{aligned}$$

となるので, Schwarz の不等式から

$$\begin{aligned} \left\{ \left[\frac{\partial}{\partial \eta} F(\xi_\theta, \eta) \right]_{\eta=\theta} \right\}^2 &= \left\{ \left[\int_{-\infty}^{\xi_\theta} \frac{\partial}{\partial \eta} f(x, \eta) d\mu(x) \right]_{\eta=\theta} \right\}^2 \\ &= \left\{ \int_{-\infty}^{\xi_\theta} \frac{\partial}{\partial \theta} f(x, \theta) d\mu(x) \right\}^2 \\ &= \left\{ \int_{-\infty}^{\xi_\theta} \frac{\frac{\partial}{\partial \theta} f(x, \theta)}{\sqrt{f(x, \theta)}} \sqrt{f(x, \theta)} d\mu(x) \right\}^2 \\ &\leq \int_{-\infty}^{\xi_\theta} \frac{\left\{ \frac{\partial}{\partial \theta} f(x, \theta) \right\}^2}{f(x, \theta)} d\mu(x) \int_{-\infty}^{\xi_\theta} f(x, \theta) d\mu(x) \\ &= \int_{-\infty}^{\xi_\theta} \frac{\left\{ \frac{\partial}{\partial \theta} f(x, \theta) \right\}^2}{f(x, \theta)} d\mu(x) \cdot F(\xi_\theta, \theta) \end{aligned}$$

となり, $F(\xi_\theta, \theta) = 1 - \alpha$ であるから

$$\left\{ \left[\frac{\partial}{\partial \eta} F(\xi_\theta, \eta) \right]_{\eta=\theta} \right\}^2 \leq (1 - \alpha) \int_{-\infty}^{\xi_\theta} \frac{\left\{ \frac{\partial}{\partial \theta} f(x, \theta) \right\}^2}{f(x, \theta)} d\mu(x) \quad (3.7)$$

となる. また, 同様にして

$$\left\{ \left[\frac{\partial}{\partial \eta} F(\xi_\theta, \eta) \right]_{\eta=\theta} \right\}^2 \leq \alpha \int_{\xi_\theta}^{\infty} \frac{\left\{ \frac{\partial}{\partial \theta} f(x, \theta) \right\}^2}{f(x, \theta)} d\mu(x) \quad (3.8)$$

となる. よって, (3.1.7), (3.1.8) より

$$\begin{aligned} I(\theta) &= \int_{-\infty}^{\infty} \frac{\left\{ \frac{\partial}{\partial \theta} f(x, \theta) \right\}^2}{f(x, \theta)} d\mu(x) \\ &= \left(\int_{-\infty}^{\xi_\theta} + \int_{\xi_\theta}^{\infty} \right) \frac{\left\{ \frac{\partial}{\partial \theta} f(x, \theta) \right\}^2}{f(x, \theta)} d\mu(x) \\ &\geq \left(\frac{1}{1 - \alpha} + \frac{1}{\alpha} \right) \left\{ \left[\frac{\partial}{\partial \eta} F(\xi_\theta, \eta) \right]_{\eta=\theta} \right\}^2 \\ &= \frac{1}{\alpha(1 - \alpha)} \left\{ \left[\frac{\partial}{\partial \eta} F(\xi_\theta, \eta) \right]_{\eta=\theta} \right\}^2 \end{aligned} \quad (3.9)$$

となるので, (3.1.6), (3.1.9) より, n が十分大きいとき

$$V_\theta(F(\hat{\xi}_\theta, \theta)) = \frac{1}{nI(\theta)} \left\{ \left[\frac{\partial}{\partial \eta} F(\xi_\theta, \eta) \right]_{\eta=\theta} \right\}^2 + o\left(\frac{1}{n}\right) \leq \frac{\alpha(1 - \alpha)}{n} \quad (3.10)$$

が成り立つ.

一般に, 確率変数 X_1, \dots, X_n がたがいに独立にいずれも c.d.f. $F(x)$ をもつ分布に従うとすると, 経験分布関数 (e.d.f.) $\hat{F}_n(x) := (1/n)\#\{i \mid X_i \leq x\}$ を考える. 各 $i = 1, 2, \dots, n$ について

$$Y_i = \begin{cases} 1 & (X_i \leq x), \\ 0 & (X_i > x) \end{cases}$$

とすると, Y_1, \dots, Y_n はたがいに独立に $\text{Ber}(F(x))$ に従うので, $\sum_{i=1}^n Y_i$ は 2 項分布 $B(n, F(x))$ に従う. また $\bar{Y} := (1/n) \sum_{i=1}^n Y_i = \hat{F}_n(x)$ より, その分散は

$$V(\bar{Y}) = V(\hat{F}_n(x)) = \frac{1}{n} V(Y_1) = \frac{1}{n} F(x)(1 - F(x))$$

になる. ここで, c.d.f. $F(x)$ を $F(x, \theta)$ とすれば

$$F(\xi_\theta, \theta) = 1 - \alpha$$

であるとき

$$F(\xi_\theta, \theta) \left(1 - F(\xi_\theta, \theta) \right) = (1 - \alpha)\alpha$$

となるから,

$$V_{\theta}(\hat{F}_n(\xi_{\theta})) = \frac{\alpha(1-\alpha)}{n}$$

となり, \hat{F}_n の分散が上界と一致し, 最悪な場合となる.

3.2 ノンパラメトリックの場合

観測可能な確率ベクトルを $\mathbf{X} := (X_1, \dots, X_n)$, 未観測の確率変数を Y とし, X_1, \dots, X_n, Y をたがいに独立に同じ連続型分布 F に従うとする. ここで X_1, \dots, X_n, Y の順序統計量を $O_{n+1} = \{Z_1 < \dots < Z_n < Z_{n+1}\}$ とすると

$$P\{Y = Z_i | O_{n+1}\} = \frac{1}{n+1} \quad (i = 1, \dots, n)$$

となる. このとき $j := (n+1)\alpha$ が整数であると仮定すると

$$P\{Y \leq Z_{n+1-j} | O_{n+1}\} = \frac{n+1-j}{n+1} = 1 - \frac{j}{n+1} = 1 - \alpha$$

となる. また X_1, \dots, X_n の順序統計量を $X_{(1)} \leq \dots \leq X_{(n)}$ とすると

$$Y \leq Z_{n+1-j} \Leftrightarrow Y \leq X_{(n+1-j)}$$

となるから

$$P\{Y \leq X_{(n+1-j)} | X_{(1)} \leq \dots \leq X_{(n)}\} = 1 - \alpha$$

となる. ここで, F は単調増加であるから

$$P\{F(Y) \leq F(X_{(n+1-j)}) | X_{(1)} \leq \dots \leq X_{(n)}\} = 1 - \alpha$$

となる. いま

$$U_{(i)} := F(X_{(i)}), \quad V := F(Y)$$

とすれば, $U_{(1)} < \dots < U_{(n)}$ は一様分布 $U(0, 1)$ からの順序統計量となり, V は一様分布 $U(0, 1)$ に従う. また

$$Y \leq X_{(n+1-j)} \Leftrightarrow V \leq U_{(n+1-j)}$$

であるから

$$1 - \alpha = P\{V \leq U_{(n+1-j)} | U_{(1)} < \dots < U_{(n)}\} = U_{(n+1-j)}$$

になる. このとき $T := U_{(n+1-j)}$ とおけば, T の p.d.f. は

$$f(t) = \frac{n!}{(n-j)!(j-1)!} t^{n-j} (1-t)^{j-1} \quad (0 \leq t \leq 1)$$

となる. このとき

$$E(T) = 1 - \frac{j}{n+1}, \quad E(T^2) = \frac{(n+2-j)(n+1-j)}{n^2 + 3n + 2}$$

になり, $j = (n+1)\alpha$ であるから

$$V(U_{(n+1-j)}) = V(T) = E(T^2) - \{E(T)\}^2 = \frac{\alpha(1-\alpha)}{n+2}$$

となり, $\alpha(1-\alpha)/n$ より小さくなり, その差は表 3.1 のようになる.

$n \setminus 1 - \alpha$	0.95	0.90	0.80	0.50
10	79.1667	150.0000	266.6667	416.6667
20	21.5909	40.9091	72.7273	113.6364
50	3.6538	6.9231	12.3077	19.2308
100	0.9314	1.7647	3.1373	4.9020

表 3.1 差 $\{\alpha(1 - \alpha)/n\} - \{\alpha(1 - \alpha)/(n + 2)\}$ の値 ($\times 10^{-5}$)

3.3 平均が未知の正規分布の場合

観測可能な確率ベクトルを $X := (X_1, \dots, X_n)$, 未観測の確率変数を Y とし, X_1, \dots, X_n, Y がたがいに独立にいずれも正規分布 $N(\theta, 1)$ に従うとする. ただし, θ は未知とする. このとき, $\bar{X} := (1/n) \sum_{i=1}^n X_i$ は θ の一致推定量となる ([A03]).

いま, Φ を $N(0, 1)$ の c.d.f. とすれば, $F(x, \theta) = \Phi(x - \theta)$ であり, Φ は単調増加関数であるから, $1 - \alpha = \Phi(\xi_\theta - \theta)$ より

$$\xi_\theta = \Phi^{-1}(1 - \alpha) + \theta$$

となるので

$$\hat{\xi}_\theta = \xi_{\hat{\theta}} = \Phi^{-1}(1 - \alpha) + \hat{\theta} = u_\alpha + \bar{X}$$

となる. ただし, u_α は $N(0, 1)$ の上側 $100\alpha\%$ 点とする. したがって

$$F(\hat{\xi}_\theta, \theta) = \Phi(u_\alpha + \bar{X} - \theta)$$

となる. 一方, \bar{X} は $N(\theta, 1/n)$ に従うから, $Z := \sqrt{n}(\bar{X} - \theta)$ は $N(0, 1)$ に従う. よって, ϕ を $N(0, 1)$ の p.d.f. とすれば, n が十分大きいとき

$$\begin{aligned} V(F(\hat{\xi}_\theta, \theta)) &= V\left(\Phi\left(u_\alpha + \frac{Z}{\sqrt{n}}\right)\right) \\ &= V\left(\Phi(u_\alpha) + \phi(u_\alpha) \frac{Z}{\sqrt{n}}\right) + o\left(\frac{1}{n}\right) \\ &= \frac{\{\phi(u_\alpha)\}^2}{n} V(Z) + o\left(\frac{1}{n}\right) \\ &= \frac{\{\phi(u_\alpha)\}^2}{n} + o\left(\frac{1}{n}\right) \end{aligned} \quad (3.11)$$

になる. また, $\{\phi(u_\alpha)\}^2 \leq \alpha(1 - \alpha)$ になる. 実際, $\Phi(u_\alpha) = 1 - \alpha$ より, 任意の $x \in \mathbf{R}^1$ に対して

$$\Phi(x) - \{\Phi(x)\}^2 - \{\phi(x)\}^2 \geq 0$$

を示せばよい. ここで

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

であるから

$$\phi'(x) = -x\phi(x), \quad \phi''(x) = (x^2 - 1)\phi(x)$$

になる。いま

$$f(x) := \Phi(x) - \{\Phi(x)\}^2 - \{\phi(x)\}^2$$

とおくと

$$\begin{aligned} f'(x) &= \phi(x) - 2\Phi(x)\phi(x) - 2\phi(x)\phi'(x) \\ &= \phi(x)\{1 - 2\Phi(x) - 2\phi'(x)\} \end{aligned}$$

となり、また

$$g(x) := 1 - 2\Phi(x) - 2\phi'(x)$$

とおくと、任意の $x \in \mathbf{R}^1$ について

$$g'(x) = -2\phi(x) - 2\phi''(x) = -2\phi(x) - 2(x^2 - 1)\phi(x) = -2x^2\phi(x) \leq 0$$

になる。したがって、 $g'(x) \leq 0$ ($x \in \mathbf{R}^1$) より、 $g(x)$ は非増加関数である。さらに

$$g(0) = 1 - 2\Phi(0) - 2\phi'(0) = 1 - 1 - 0 = 0$$

であるから、 $x \geq 0$ において、 $g(x) \leq 0$ となる。よって $x \geq 0$ において、 $f(x)$ は $f'(x) \leq 0$ から非増加関数である。また $f(x)$ は $x = 0$ において対称な関数であるから $x < 0$ において非減少関数である。ここで

$$f(0) = \Phi(0) - \{\Phi(0)\}^2 - \{\phi(0)\}^2 > 0.5 - 0.25 - 0.16 = 0.09 > 0$$

でかつ

$$\lim_{x \rightarrow \infty} f(x) = \lim_{x \rightarrow -\infty} f(x) = 0$$

であるから、任意の $x \in \mathbf{R}^1$ に対して

$$\Phi(x) - \{\Phi(x)\}^2 - \{\phi(x)\}^2 \geq 0$$

となる。したがって $\{\phi(u_\alpha)\}^2 \leq \alpha(1 - \alpha)$ になる。ここで、(3.1.10) の上界と (3.3.1) の差の n 倍の値については表 3.2、図 3.1 のようになる。

$1 - \alpha$	$\alpha(1 - \alpha) - \{\phi(u_\alpha)\}^2$
0.95	3.6863
0.90	5.9200
0.80	8.1621
0.50	9.0845

表 3.2 差 $\alpha(1 - \alpha) - \{\phi(u_\alpha)\}^2$ の値 ($\times 10^{-2}$)

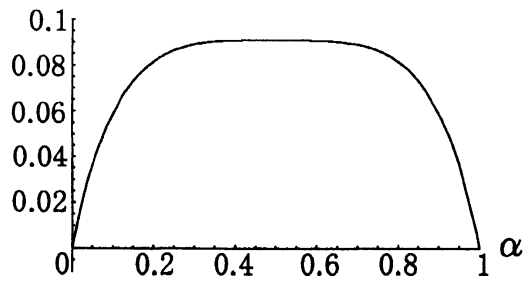


図 3.1 差 $\alpha(1-\alpha) - \{\phi(u_\alpha)\}^2$ のグラフ

3.4 分散が未知の正規分布の場合

観測可能な確率ベクトルを $X := (X_1, \dots, X_n)$, 未観測の確率変数を Y とし, X_1, \dots, X_n, Y をたがいに独立にいずれも正規分布 $N(0, \theta)$ に従うとする. ただし, θ は未知とする. このとき, $(1/n) \sum_{i=1}^n X_i^2$ は θ の一致推定量となる ([A03]).

いま, Φ を $N(0, 1)$ の c.d.f. とすれば, $F(x, \theta) = \Phi(x/\sqrt{\theta})$ であり, Φ は単調増加関数であるから, $1-\alpha = \Phi(\xi_\theta/\sqrt{\theta})$ より

$$\xi_\theta = \sqrt{\theta} \Phi^{-1}(1-\alpha)$$

となるので

$$\hat{\xi}_\theta = \xi_\theta = \sqrt{\hat{\theta}} \Phi^{-1}(1-\alpha) = u_\alpha \sqrt{\frac{1}{n} \sum_{i=1}^n X_i^2}$$

となる. ただし, u_α は $N(0, 1)$ の上側 $100\alpha\%$ 点とする. したがって

$$F(\hat{\xi}_\theta, \theta) = \Phi \left(u_\alpha \sqrt{\frac{1}{n\theta} \sum_{i=1}^n X_i^2} \right) = \Phi \left(u_\alpha \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{X_i}{\sqrt{\theta}} \right)^2} \right)$$

となる.

一方, $X_i/\sqrt{\theta}$ は $N(0, 1)$ に従うので, $(X_i/\sqrt{\theta})^2$ は自由度 1 のカイ 2 乗分布に従う. また各 i に対してたがいに独立であるから, よってカイ 2 乗分布の再生性より

$$\sum_{i=1}^n \left(\frac{X_i}{\sqrt{\theta}} \right)^2 \sim \chi_n^2$$

となる. ここで, $Z := \sqrt{\sum_{i=1}^n (X_i/\sqrt{\theta})^2} \sim \chi_n$ より

$$V(Z) = n - 2 \left(\frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \right)^2$$

となるので, Stirling の公式より $V(Z) = 1/2 + o(1)$ となる. よって, ϕ を $N(0, 1)$ の p.d.f. とすれば, n が十分大きいとき

$$\begin{aligned}
V(F(\hat{\xi}_\theta, \theta)) &= V\left(\Phi\left(u_\alpha \frac{Z}{\sqrt{n}}\right)\right) \\
&= V\left(\Phi\left(u_\alpha + u_\alpha\left(\frac{Z}{\sqrt{n}} - 1\right)\right)\right) \\
&= V\left(\Phi(u_\alpha) + u_\alpha\phi(u_\alpha)\left(\frac{Z}{\sqrt{n}} - 1\right)\right) + o\left(\frac{1}{n}\right) \\
&= V\left(\Phi(u_\alpha) + u_\alpha\phi(u_\alpha) \cdot \frac{Z}{\sqrt{n}} - u_\alpha\phi(u_\alpha)\right) + o\left(\frac{1}{n}\right) \\
&= \frac{u_\alpha^2\{\phi(u_\alpha)\}^2}{n}V(Z) + o\left(\frac{1}{n}\right) \\
&= \frac{u_\alpha^2\{\phi(u_\alpha)\}^2}{2n} + o\left(\frac{1}{n}\right) \tag{3.12}
\end{aligned}$$

となり, これは $\alpha(1-\alpha)/n$ 以下となる. 実際, $\Phi(u_\alpha) = 1-\alpha$ より, 任意の $x \in \mathbf{R}^1$ に対して

$$2\Phi(x) - 2\{\Phi(x)\}^2 - x^2\{\phi(x)\}^2 \geq 0 \tag{3.13}$$

を示せばよい. ここで,

$$\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

であるから

$$\phi'(x) = -x\phi(x), \quad \phi''(x) = (x^2 - 1)\phi(x)$$

になる. いま

$$f(x) := 2\Phi(x) - 2\{\Phi(x)\}^2 - x^2\{\phi(x)\}^2$$

とおくと

$$\begin{aligned}
f'(x) &= 2\phi(x) - 4\Phi(x)\phi(x) - 2x\{\phi(x)\}^2 - 2x^2\phi(x)\phi'(x) \\
&= 2\phi(x)\{1 - 2\Phi(x) - x\phi(x) - x^2\phi'(x)\}
\end{aligned}$$

となり, また

$$g(x) := 1 - 2\Phi(x) - x\phi(x) - x^2\phi'(x)$$

とおき, $x \geq 0$ について $g(x)$ を考える. ここで

$$\begin{aligned}
g'(x) &= -\phi(x) - x\phi'(x) - 2x\phi'(x) - 2\phi(x) - x^2\phi''(x) \\
&= -\phi(x)(x^4 - 4x^2 + 3)
\end{aligned}$$

となり, $x \geq 0$ において $g(x)$ は $x = 1, \sqrt{3}$ で極値をとる. このとき極小値が $g(1) \doteq -0.68$, 極大値が $g(\sqrt{3}) \doteq -0.61$, $g(0) = 0$ であるから, $x \geq 0$ において, $g(x) \leq 0$ となる. よって

$x \geq 0$ において, $f(x)$ は $f'(x) \leq 0$ から非増加関数である. また $f(x)$ は $x = 0$ において対称な関数であるから $x < 0$ において非減少関数である. ここで

$$f(0) = 2\Phi(0) - 2\{\Phi(0)\}^2 - \{\phi(0)\}^2 \cdot 0 = 1 - 0.5 = 0.5 > 0$$

でかつ

$$\lim_{x \rightarrow \infty} f(x) = \lim_{x \rightarrow -\infty} f(x) = 0$$

であるから, 任意の $x \in \mathbf{R}^1$ に対して

$$2\Phi(x) - 2\{\Phi(x)\}^2 - \{\phi(x)\}^2 x^2 \geq 0$$

となり, (3.4.2) が示された. したがって $\alpha(1-\alpha) - u_\alpha^2 \{\phi(u_\alpha)\}^2 / 2 \geq 0$ になり, (3.1.10) の上界と (3.4.1) の差の n 倍の値については表 3.3, 図 3.2 のようになる.

$1 - \alpha$	$\alpha(1 - \alpha) - u_\alpha^2 \{\phi(u_\alpha)\}^2 / 2$
0.95	3.3111
0.90	6.4708
0.80	13.2241
0.50	25.0000

表 3.3 差 $\alpha(1 - \alpha) - [u_\alpha^2 \{\phi(u_\alpha)\}^2 / 2]$ の値 ($\times 10^{-2}$)

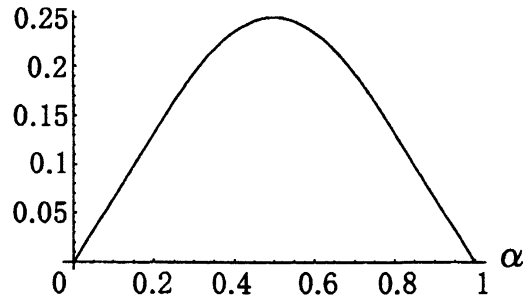


図 3.2 差 $\alpha(1 - \alpha) - [u_\alpha^2 \{\phi(u_\alpha)\}^2 / 2]$ のグラフ

3.5 母数が未知の指数分布の場合

観測可能な確率ベクトルを $\mathbf{X} := (X_1, \dots, X_n)$, 未観測の確率変数を Y とし, X_1, \dots, X_n, Y をたがいに独立にいずれも指数分布 $\text{Exp}(\lambda)$ に従うとする. ただし, λ は未知とする. このとき, p.d.f., c.d.f. はそれぞれ

$$f(x, \lambda) = \frac{1}{\lambda} e^{-x/\lambda}, \quad F(x, \lambda) = 1 - e^{-x/\lambda} \quad (x > 0)$$

であり, \bar{X} は λ の一致推定量となる ([A03]).

ここで, 各 $i = 1, \dots, n$ について $Z_i := X_i / \lambda$ とおくと, Z_1, \dots, Z_n はたがいに独立にいずれも指数分布 $\text{Exp}(1)$, すなわちガンマ分布 $G(1, 1)$ に従い, その c.d.f. は

$$F_0(z) = \begin{cases} 1 - e^{-z} & (z > 0), \\ 0 & (z \leq 0) \end{cases}$$

となる。ここで、ガンマ分布の再生性から

$$\sum_{i=1}^n Z_i = \frac{1}{\lambda} \sum_{i=1}^n X_i$$

はガンマ分布 $G(n, 1)$ に従う。したがって、いま ξ_λ を $\text{Exp}(\lambda)$ の上側 $100\alpha\%$ 点とすると $F(x, \lambda) = F_0(x/\lambda)$ より $1 - \alpha = F_0(\xi_\lambda/\lambda)$ となり、 F_0 は $x > 0$ において単調増加関数であるから

$$\xi_\lambda = \lambda F_0^{-1}(1 - \alpha)$$

になる。このとき、 λ の有効推定量 $\hat{\lambda} = \bar{X}$ を用いて ξ_λ の推定量

$$\hat{\xi}_\lambda = \xi_\lambda = \hat{\lambda} F_0^{-1}(1 - \alpha) = \bar{X} F_0^{-1}(1 - \alpha)$$

を考える。いま、指数分布 $\text{Exp}(1)$ の上側 $100\alpha\%$ 点を k_α とすれば $k_\alpha = F_0^{-1}(1 - \alpha)$ となり、 $\hat{\xi}_\lambda = k_\alpha \bar{X}$ となるので、 n が十分大きいとき、 f_0 を $\text{Exp}(1)$ の p.d.f. とすれば

$$\begin{aligned} F_0\left(\frac{\hat{\xi}_\lambda}{\lambda}\right) &= F_0\left(k_\alpha \frac{\bar{X}}{\lambda}\right) = F_0\left(k_\alpha + k_\alpha \left(\frac{\bar{X}}{\lambda} - 1\right)\right) \\ &= F_0(k_\alpha) + k_\alpha f_0(k_\alpha) \left(\frac{\bar{X}}{\lambda} - 1\right) + o_p\left(\frac{1}{\sqrt{n}}\right) \\ &= F_0(k_\alpha) + \left(\frac{k_\alpha \sum_{i=1}^n X_i}{n\lambda} - k_\alpha\right) f_0(k_\alpha) + o_p\left(\frac{1}{\sqrt{n}}\right) \\ &= F_0(k_\alpha) + \left(\frac{k_\alpha \sum_{i=1}^n Z_i}{n}\right) f_0(k_\alpha) - k_\alpha f_0(k_\alpha) + o_p\left(\frac{1}{\sqrt{n}}\right) \end{aligned}$$

となる。したがって

$$\begin{aligned} V(F(\hat{\xi}_\lambda, \lambda)) &= V\left(\left(\frac{k_\alpha \sum_{i=1}^n Z_i}{n}\right) f_0(k_\alpha)\right) + o\left(\frac{1}{n}\right) \\ &= \frac{k_\alpha^2 \{f_0(k_\alpha)\}^2}{n^2} V\left(\sum_{i=1}^n Z_i\right) + o\left(\frac{1}{n}\right) \\ &= \frac{k_\alpha^2 \{f_0(k_\alpha)\}^2}{n} + o\left(\frac{1}{n}\right) \end{aligned} \quad (3.14)$$

となり、これは $\alpha(1 - \alpha)/n$ 以下となる。実際、 $F_0(k_\alpha) = 1 - \alpha$ より、任意の $x \geq 0$ に対して

$$F_0(x)(1 - F_0(x)) - x^2 \{f_0(x)\}^2 \geq 0 \quad (3.15)$$

を示せばよい。いま、

$$f_0(x) = e^{-x}, \quad F_0(x) = 1 - e^{-x}$$

であるから

$$F_0(x)(1 - F_0(x)) - x^2 \{f_0(x)\}^2 = e^{-x}(1 - e^{-x}(1 + x^2)) \quad (3.16)$$

となる。ここで

$$g(x) := e^{-x}(1+x^2)$$

とおくと、 $x \geq 0$ に対して

$$g'(x) = 2xe^{-x} - (1+x^2)e^{-x} = -(x-1)^2e^{-x} \leq 0$$

になる。よって、 $x \geq 0$ において $g(x)$ は非単調増加関数となる。さらに

$$g(0) := e^0(1+0^2) = 1$$

であるから、 $x \geq 0$ において、 $g(x) \leq 1$ となる。したがって、 $x \geq 0$ について

$$e^{-x}(1 - e^{-x}(1+x^2)) \geq 0$$

になり、(3.5.3) から (3.5.2) は示された。ここで (3.1.10) の上界と (3.5.1) の差の n 倍の値については表 3.4、図 3.3 のようになる。また、 β が既知のワイブル分布 $W(\beta, \lambda)$ もこの形に帰着する。

$1 - \alpha$	$\alpha(1 - \alpha) - k_\alpha^2 \{f_0(k_\alpha)\}^2$
0.95	4.5126
0.90	8.1008
0.80	12.8132
0.50	12.9887

表 3.4 差 $\alpha(1 - \alpha) - [k_\alpha^2 \{f_0(k_\alpha)\}^2]$ の値 ($\times 10^{-2}$)

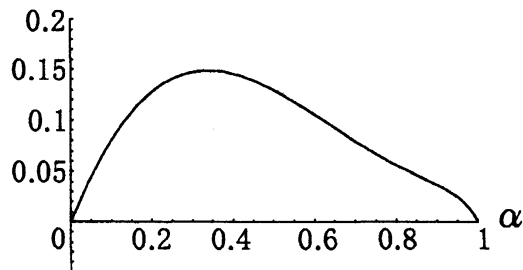


図 3.3 差 $\alpha(1 - \alpha) - [k_\alpha^2 \{f_0(k_\alpha)\}^2]$ のグラフ

3.6 パラメトリックの場合とノンパラメトリックの場合との比較

パラメトリックの場合とノンパラメトリックの場合において予測限界の分布関数の分散の上界の比較を行う。つまり、第3.2節よりノンパラメトリックの場合には予測限界の分布関数の分散の上界が $\alpha(1-\alpha)/(n+2)$ となるから、分布型を変えたときの予測限界の分布関数の分散の上界とそれの差を比較する。

まず、平均が未知の正規分布の場合には、第3.3節より予測限界の分布関数の分散が $\{\phi(u_\alpha)\}^2/n$ であるから、それらの差の値は表3.5と図3.4のようになる。

$n \setminus 1 - \alpha$	0.95	0.90	0.80	0.50
10	28.9464	44.2003	54.9547	49.1784
20	16.2724	25.5093	33.5379	34.0589
50	7.0072	11.1478	15.0935	16.2459
100	3.5932	5.7436	7.8484	8.5943

表3.5 差 $[\alpha(1-\alpha)/(n+2)] - [\{\phi(u_\alpha)\}^2/n]$ の値 ($\times 10^{-4}$)

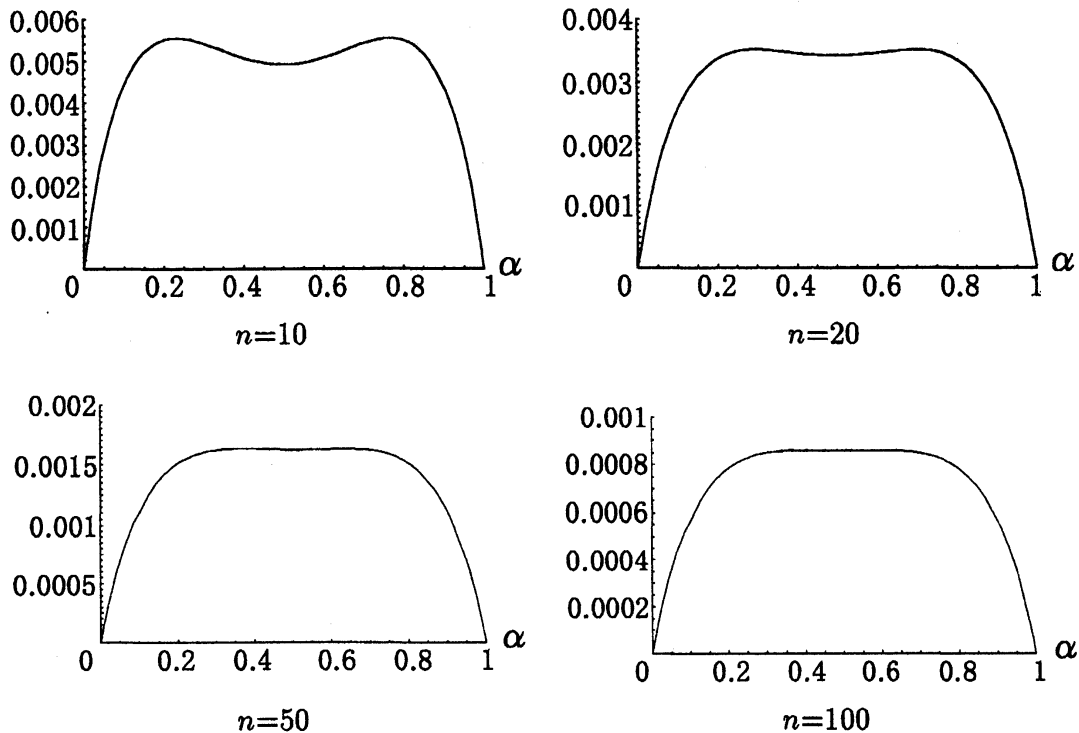


図3.4 差 $[\alpha(1-\alpha)/(n+2)] - [\{\phi(u_\alpha)\}^2/n]$ のグラフ

表3.5と図3.4より明らかに平均が未知の正規分布を仮定したほうがノンパラメトリックの場合より精度が上がっている。また、 n が小さいときには α が0.2から0.25の間で最大になることがわかる。

次に、分散が未知の正規分布の場合には、第3.4節より予測限界の分布関数の分散が $u_\alpha^2 \{\phi(u_\alpha)\}^2 / 2n$ であるから、それらの差の値は表3.6と図3.5のようになる。

$n \setminus 1 - \alpha$	0.95	0.90	0.80	0.50
10	25.5703	50.3692	106.3005	208.3333
20	14.4883	28.4248	59.0255	113.6364
50	6.2713	12.2748	25.2455	48.0769
100	3.2215	6.3007	12.9174	24.5098

表3.6 差 $[\alpha(1-\alpha)/(n+2)] - [u_\alpha^2 \{\phi(u_\alpha)\}^2 / 2n]$ の値 ($\times 10^{-4}$)

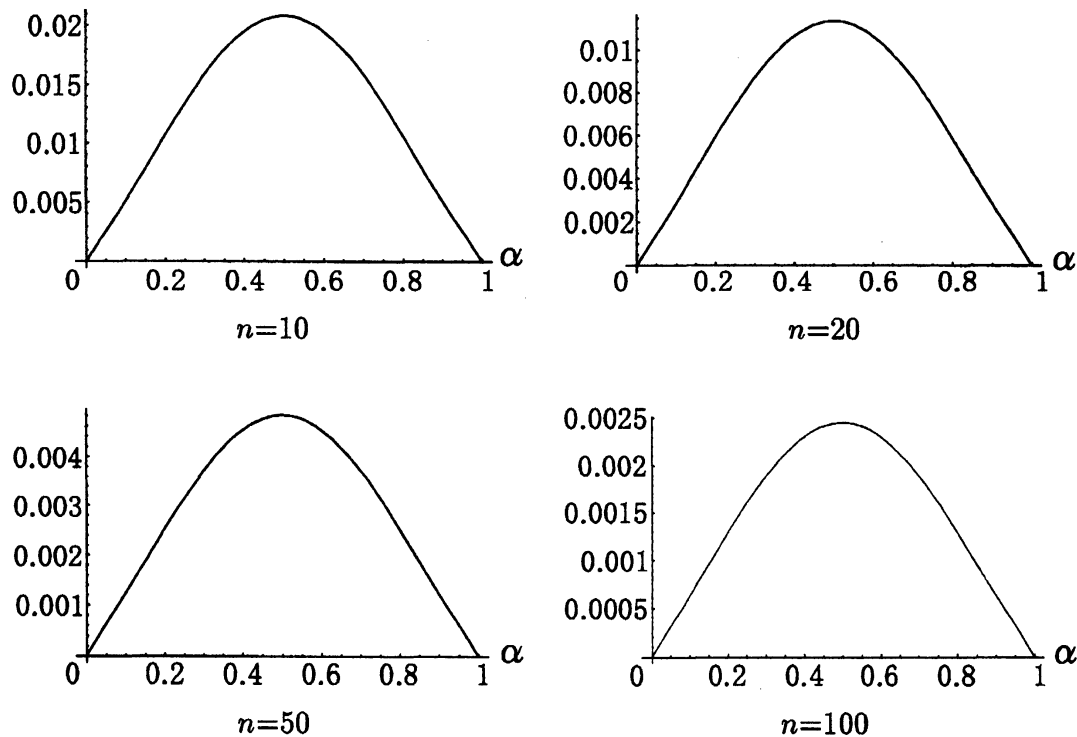


図3.5 差 $[\alpha(1-\alpha)/(n+2)] - [u_\alpha^2 \{\phi(u_\alpha)\}^2 / 2n]$ のグラフ

この場合も同様に、表3.6と図3.5から分散が未知の正規分布を仮定したほうがノンパラメトリックの場合より精度が上がっている。また、 α が0.5に近づけば近づくほど、精度が上がっていることがわかる。

最後に、母数が未知の指数分布の場合には、第3.5節より予測限界の分布関数の分散が $k_\alpha^2 \{f_0(k_\alpha)\}^2/n$ であるから、それらの差の値は表3.7と図3.6のようになる。

$n \setminus 1 - \alpha$	0.95	0.90	0.80	0.50
10	37.2089	66.0083	101.4658	88.2201
20	20.4037	36.4133	56.7935	53.5797
50	8.6597	15.5094	24.3957	24.0543
100	4.4194	7.9244	12.4995	12.4985

表3.7 差 $[\alpha(1-\alpha)/(n+2)] - [k_\alpha^2 \{f_0(k_\alpha)\}^2/n]$ の値 ($\times 10^{-4}$)

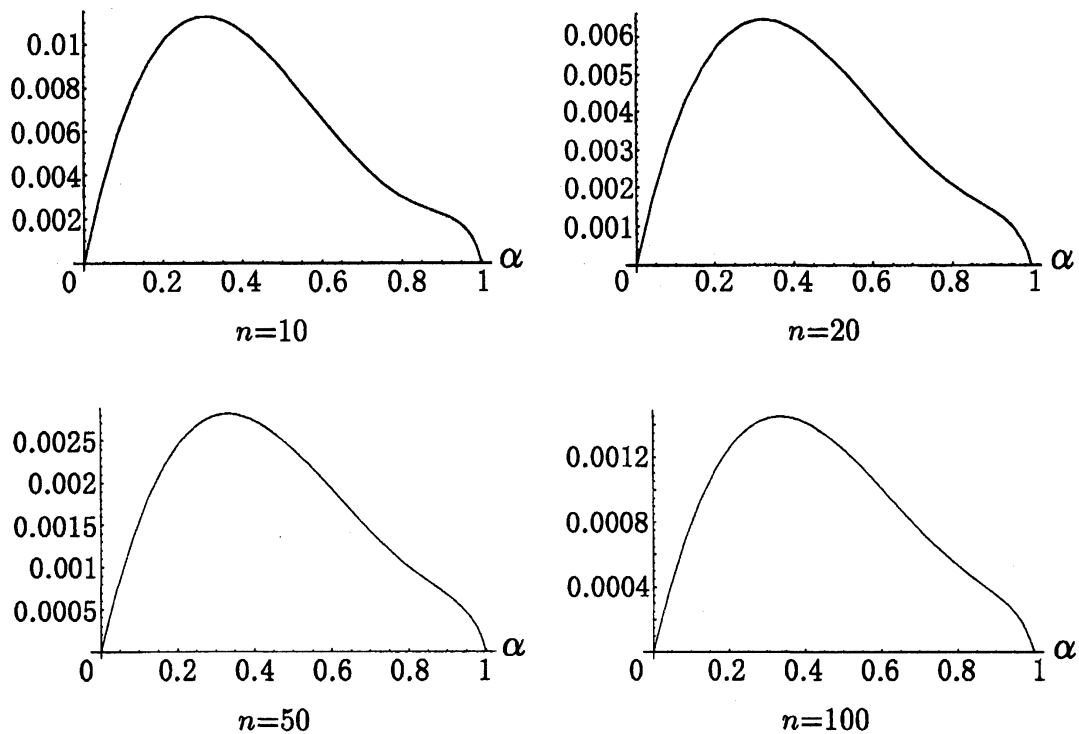


図3.6 差 $[\alpha(1-\alpha)/(n+2)] - [k_\alpha^2 \{f_0(k_\alpha)\}^2/n]$ のグラフ

この場合も同様に、表3.7と図3.6から母数が未知の指数分布を仮定したほうがノンパラメトリックの場合より精度が上がっている。また、 α が0.3あたりで最大になっていることがわかる。

4 おわりに

本論では、まず、予測尤度の性質、またそれに基づく予測域の構成について述べ、次に未観測の確率変数の予測限界を累積分布関数を通して分散によって評価し、パラメトリックモデルとノンパラメトリックモデルの場合の差異について考察した。

実際に、未観測の確率変数の上側予測限界による下側確率の分散の上界を求め、ノンパラメトリックのときに順序統計量を用いる場合や既知の分布型が分かっている場合の予測限界の分散と比較した。また、パラメトリックの場合とノンパラメトリックの場合との比較も行った。その結果、パラメトリックの場合の方が予測限界の分散が小さくなることを示すとともに、予測の精度の度合を数値的に示した。そして、いずれの場合も、未観測の確率変数の上側予測限界による下側確率の分散の上界は比較的良い上界であることも確認された。本論では、予測限界の評価の一つの方法を提案したが、この他にも予測域の被覆確率等による評価法も考えられ、今後の課題になる。

参考文献

- [A03] 赤平昌文 (2003). 統計解析入門. 森北出版.
- [B90] Bjørnstad, J. F. (1990). Predictive Likelihood : A Review. *Statistical Science* **5** (1), 242-265.
- [G93] Geisser, S. (1993). *Predictive Inference : An Introduction*. Chapman & Hall, New York.
- [HH98] Howlader, H. A., Hossain, A. (1998). Bayesian prediction intervals for Maxwell parameters. *Metron* **56**, 153-168.
- [H00] 飛田英祐 (2000). Non-Bayesian and Bayesian prediction intervals for an unobservable chi random variable. 京都大学 数理解析研究所講究録 **1161**, 1-11.
- [T75] 竹内啓 (1975). 統計的予測論. 培風館.