

問題の概要

いつもの指数分布。今回は観測が少し変な状況で、元の分布のパラメータを最尤推定できるかという問題。指数分布の無記憶性を覚えていると、(3) は思いつきやすいかもしれない。(4) は式を整理して、いくつかの計算テクニックを使う必要がある。難しい問題だが、教育的な問題なので是非解いてみてほしい。

(1)

$$\begin{aligned} g(\mu) &= \int_0^a af(x; \mu) dx + \int_a^\infty xf(x; \mu) dx \\ &= a + \mu e^{-\frac{a}{\mu}} \end{aligned}$$

(2)

$g'(\mu) = \left(1 + \frac{a}{\mu}\right) \exp\left(-\frac{a}{\mu}\right) > 0$ と $g(0) = a$ より、 $g(\hat{\mu}) = \bar{Y} > a$ なる $\hat{\mu}$ がただ一つ存在する。

(3)

かなり難しい。

まず、指数分布の和の分布はアーラン分布(ガンマ分布)になることを簡単に示す。 X_1, X_2 が平均 μ の指数分布に従うとすると、 $X_1 + X_2$ の密度関数は畳み込み積分により、 $z\mu^{-2} \exp\left(-\frac{z}{\mu}\right)$ となる。これを使って $X_1 + X_2 + X_3$ の密度関数は畳み込み積分により、 $\frac{1}{2}z^2\mu^{-3} \exp\left(-\frac{z}{\mu}\right)$ となる。これを一般化して、 $X_1 + \dots + X_n$ の分布は $\frac{z^{n-1}}{(n-1)!} \mu^{-n} \exp\left(-\frac{z}{\mu}\right)$ となる。詳しくは 2024 年問 4 (3) を見てほしい、別解も載っている。この分布の確率密度関数を $f(x; n, \lambda)$ とする。

指数分布の無記憶性 $P(X > b \mid X > a) = P(X > b - a)$ は簡単に示せる。準備が終わったので、早速示していく。

$$P(M = m, \bar{Y} \leq b) = \binom{n}{m} P(Y = a)^m P(Y > a)^{n-m} P\left(\sum_{i=1}^{n-m} Y_i \leq nb - ma \mid Y_i > a\right)$$

ここで指数分布の無記憶性を各 Y_i に用いて

$$\begin{aligned} &= \dots P\left(\sum_{i=1}^{n-m} Y_i \leq n(b-a)\right) \\ &= \dots \int_0^{n(b-a)} f(x; n-m, \mu) dx \end{aligned}$$

ここで $Y = \frac{x}{n} + a$ において積分区間を調整すると

$$= \dots \int_a^b nf(n(y-a), n-m, \mu) dy$$

よって

$$h(m, y; \mu) = \binom{n}{m} \left(1 - e^{-\frac{a}{\mu}}\right)^m \left(e^{-\frac{a}{\mu}}\right)^{n-m} \frac{(y-a)^{n-m-1}}{(n-m-1)! \left(\frac{\mu}{n}\right)^{n-m}} e^{-\frac{n(y-a)}{\mu}}$$

となる。指数部分を整理すると、

$$h(m, y; \mu) = \binom{n}{m} \left(e^{\frac{a}{\mu}} - 1\right)^m \frac{(y-a)^{n-m-1}}{(n-m-1)! \left(\frac{\mu}{n}\right)^{n-m}} e^{-\frac{ny}{\mu}} \quad (1)$$

(4)

順当にやれば、 $h(m, y; \mu)$ を μ で微分して 0 となる点が 1 つであればよい。しかし、これはかなり難しい。

🔗 テクニック

確率密度の極値を探すとき、対数を取ってから微分すると楽な場合が多い。確率密度は非負かつ積の形になっている傾向があるため効果的である。また、解析的に零点を求めるのが難しい場合は、更に微分することで零点の存在を示せる場合がある。(1) はその誘導であると考えられる。また、指数分布などはパラメータの表現を、平均を表す μ と逆数の λ を使い分けることで、特定の計算が楽になる場合がある。今回は $\lambda = \frac{1}{\mu}$ とおくことで、微分が簡単になる。

まず、 $\lambda := \frac{1}{\mu} > 0$ と置いて、式 1 の対数を取ると、

$$\log h(m, y; \mu) = C(m, y) + m \log(e^{a\lambda} - 1) + (n-m) \log \lambda - ny\lambda$$

λ で微分すると、

$$\frac{d}{d\lambda} \log h = \frac{mae^{a\lambda}}{e^{a\lambda} - 1} + \frac{n-m}{\lambda} - ny$$

これの零点が唯一であることを示せばよい。解析的には求められそうにない。もう一度微分すると、

$$\frac{d^2}{d\lambda^2} \log h = -\frac{ma^2 e^{a\lambda}}{(e^{a\lambda} - 1)^2} - \frac{n-m}{\lambda^2} < 0$$

よって $(\log h)'$ は単調減少。また、 $n > m, y > a$ に注意して、

$$\begin{aligned} \lim_{\lambda \rightarrow +0} \frac{d}{d\lambda} \log h &= \infty \\ \lim_{\lambda \rightarrow \infty} \frac{d}{d\lambda} \log h &= ma - ny < m(a - y) < 0 \end{aligned}$$

より、 λ に関する方程式 $\frac{d}{d\lambda} \log h = \frac{h'}{h} = 0$ は $\lambda > 0$ でただ一つの解 λ^* を持ち、 $\lambda < \lambda^*$ では $h' > 0$ より h は増加し、 $\lambda > \lambda^*$ では $h' < 0$ より h は減少する。したがって、 h は λ^* で最大値をとる。