

問題の概要

指数分布の問題。最後の問題はポアソン分布を知っていれば理解しやすいだろう。ポアソン分布は、間隔が指数分布に従う事象が単位時間以内に起こる数 N の分布である。(2) が詰まりやすいかもしれない。少し計算量が多いが、やるべきことは単純。

(1)

$P(X_1 - c \mid X_1 \geq c)$ を求めたい。 $P(X_1 \geq c) = \exp(-\lambda c)$ なので、

$$\begin{aligned} P(X_1 - c = y \mid X_1 \geq c) &= \frac{P(X_1 = y + c \wedge X_1 \geq c)}{P(X_1 \geq c)} = \frac{\lambda \exp(-\lambda(y + c))}{\exp(-\lambda c)} \\ &= \lambda \exp(-\lambda y) \end{aligned}$$

(2)

よくあるテクニックだが、 \min や \max の分布は累積分布関数を考えるのが良い。 $Y = \min(X_1, X_2)$ について、

$$\begin{aligned} P(Y \geq c) &= P(X_1 \geq c \wedge X_2 \geq c) = P(X_1 \geq c)P(X_2 \geq c) \\ &= \exp(-2\lambda c) \end{aligned}$$

これはパラメータ 2λ の指数分布の累積分布関数なので、 $P(Y = y) = 2\lambda \exp(-2\lambda y)$ である。累積分布関数を微分して出してもよい。

$Z - Y = |X_1 - X_2|$ であることに着目すると、

$$\begin{aligned} P(Z - Y = w) &= \int_0^\infty P(X_1 = x)P(X_2 = w + x) dx + \int_0^\infty P(X_2 = x)P(X_1 = w + x) dx \\ &= 2 \int_0^\infty P(X_1 = x)P(X_2 = w + x) dx \\ &= 2\lambda^2 \exp(-\lambda w) \int_0^\infty \exp(-2\lambda x) dx \\ &= \lambda \exp(-\lambda w) \end{aligned} \tag{1}$$

である。最初の行の変形は $X_1 - X_2 = W, X_2 = X$ と変数変換してヤコビアンを計算すると 1 になるので正当化される。畳み込み分布の計算に似ているが異なるので、解答の時は一応ヤコビアンを考えて変形が正しいことを確認したほうが良いだろう。つまり、

$$P(|X_1 - X_2| = w) = P(X_1 - X_2 = w) + P(X_2 - X_1 = w)$$

としたあと、最初の項は $W = X_1 - X_2, X = X_2$ と変数変換、後ろの項は $W = X_2 - X_1, X = X_1$ として変数変換すればどちらもヤコビアンが 1 になるので、それぞれ X で積分すると最初の行の変形になる。

(3)

i.i.d. な指数分布の和の分布はアーラン分布(ガンマ分布の形状パラメータを整数に限定したもの)となる。これは畳み込み積分を使って $k = 2, 3$ ぐらいまで出してから一般化して数学的帰納法で示せばよい。

$X_1 + X_2$ の分布は、

$$\begin{aligned}P(X_1 + X_2 = z) &= \int_0^z P(X_1 = x)P(X_2 = z - x) \, dx \\&= \int_0^z \lambda^2 \exp(-\lambda z) \, dx \\&= z\lambda^2 \exp(-\lambda z)\end{aligned}$$

$X_1 + X_2 + X_3$ の分布は、

$$\begin{aligned}P(X_1 + X_2 + X_3 = z) &= \int_0^z P(X_1 + X_2 = x)P(X_3 = z - x) \, dx \\&= \int_0^z x\lambda^3 \exp(-\lambda z) \, dx \\&= \frac{1}{2}z^2\lambda^3 \exp(-\lambda z)\end{aligned}$$

もう一般系の予想はつくだろう。

$$P(X_1 + \cdots + X_n = x) = \frac{x^{n-1}}{(n-1)!} \lambda^n \exp(-\lambda x)$$

である。数学的帰納法で示す部分は簡単なので割愛する。解答では書かなくても問題ないかもしれない。

(3) 別解

畳み込み分布で示さず、特性関数を使って示す方法もある。まず、平均 $\mu = \frac{1}{\lambda}$ の指数分布の特性関数は簡単に計算でき、

$$\begin{aligned}\varphi(u) &= \int_0^\infty e^{iux} f(x; \mu) \, dx \\&= \frac{1}{\mu} \int_0^\infty e^{-\left(\frac{1}{\mu} - iu\right)x} \, dx \\&= \frac{1}{-i\mu u + 1}\end{aligned}$$

である。i.i.d.な指数分布 k 個の和 $X := X_1 + \cdots + X_k$ の特性関数 Φ は単に k 乗して、

$$\Phi(u) = \frac{1}{(-i\mu u + 1)^k}$$

ここから逆に X の密度関数 g を求める。逆フーリエ変換で、

$$g(x; \mu) = \frac{1}{2\pi} \int_{-\infty}^\infty \frac{e^{-iux}}{(-i\mu u + 1)^k} \, du$$

と書ける。 $z = -iu$ と置換して、 $x > 0$ のときを考えると、複素積分の経路を下弦の月のような、半径 R の左半円 C_R を時計周りに回る経路で取ると、

$$g(x; \mu) = \frac{i}{2\pi\mu^k} \lim_{R \rightarrow \infty} \oint_{C_R} \frac{e^{zx}}{\left(z + \frac{1}{\mu}\right)^k} dz \quad (x \geq 0)$$

非積分関数は、 $z = -\frac{1}{\mu}$ で k 位の極を持つ。これは C_R 内に入っている。留数は、

$$\begin{aligned} \frac{e^{zx}}{\left(z + \frac{1}{\mu}\right)^k} &= e^{-\frac{x}{\mu}} \frac{e^{\left(z + \frac{1}{\mu}\right)x}}{\left(z + \frac{1}{\mu}\right)^k} \\ &= e^{-\frac{x}{\mu}} \sum_{n \geq 0} \frac{x^n}{n!} \left(z + \frac{1}{\mu}\right)^{n-k} \end{aligned}$$

なので、 $n = k - 1$ の部分を見て留数がわかる。留数定理から、(反時計回りではないので -1 倍されるのに注意して)

$$\begin{aligned} g(x; \mu) &= \frac{i}{2\pi\mu^k} (-2\pi i) e^{-\frac{x}{\mu}} \frac{x^{k-1}}{(k-1)!} \\ &= e^{-\frac{x}{\mu}} \frac{x^{k-1}}{\mu^k (k-1)!} \quad (x \geq 0) \end{aligned}$$

$x < 0$ の場合は上弦の月、右半円に積分経路をとると、極を含まないので 0 とわかる。

これで、 $g(x; \mu)$ が求まった。これはガンマ分布 $G(k, \mu)$ の確率密度関数である。

(4)

N が確率変数。最尤推定量なので、 $\hat{\lambda}(N) = \arg \max_{\lambda} P(N = n \mid \lambda, a)$ である。ここで、

$$\begin{aligned} P(N = n \mid \lambda, a) &= P(X_1 + \cdots + X_N \leq a \wedge X_1 + \cdots + X_{N+1} > a) \\ &= \int_0^a P(X_1 + \cdots + X_N = c) P(X_{n+1} > a - c) dc \end{aligned}$$

2 行目の変形は(2)と同様に、 $X = X_1 + \cdots + X_N, Y = X_{N+1}$ としたときの変数変換のヤコビアンが 1 であることで正当化される。ちゃんと言うなら長くなるが今回は省略。あとは(3)を使って具体的に計算すると、

$$\begin{aligned} P(N = n \mid \lambda, a) &= \int_0^a \lambda^n \exp(-\lambda c) \frac{c^{n-1}}{(n-1)!} \exp(-\lambda(a-c)) \\ &= \lambda^n \frac{a^n}{n!} \exp(-\lambda a) \end{aligned}$$

これを λ で微分して尤度最大の λ は $\hat{\lambda} = \frac{N}{a}$ である。

さて、期待値は

$$\begin{aligned} E[\hat{\lambda}(N)] &= \sum_{n=0}^{\infty} \hat{\lambda}(N) P(N = n \mid \lambda, a) \\ &= \lambda \sum_{n=0}^{\infty} \lambda^n \frac{a^n}{n!} \exp(-\lambda a) \\ &= \lambda \sum_{n=0}^{\infty} P(N = n \mid \lambda, a) = \lambda \end{aligned}$$

となる。二次モーメントは

$$\begin{aligned} E[\hat{\lambda}(N)^2] &= \sum_{n=0}^{\infty} \frac{n^2}{a^2} P(N = n | \lambda, a) \\ &= \frac{\lambda}{a} \sum_{k=0}^{\infty} (k+1) \lambda^k \frac{a^k}{k!} \exp(-\lambda a) \\ &= \frac{\lambda}{a} (1 + E[N]) = \frac{\lambda}{a} + \lambda E\left[\frac{N}{a}\right] \\ &= \frac{\lambda}{a} + \lambda E[\hat{\lambda}(N)] = \frac{\lambda}{a} + \lambda^2 \end{aligned}$$

よって分散は

$$\frac{\lambda}{a}$$

これらの結果は直感的か？ まず、 $\hat{\lambda} = \frac{N}{a}$ はかなり直感的だろう。 N 個でだいたい a になるので平均の逆数は $\frac{N}{a}$ と考えるのは妥当だ。そして期待値が真のパラメータ λ に一致する。バイアスが 0 であることを意味する。分散は $\frac{\lambda}{a}$ だが、これは a が大きく λ が小さいほど小さくなる。これも直感的である。