

# ポアソン過程—計数分布と到着時間間隔分布—

1/08/2011

熊谷 朝臣

以前、「なぜ降雨イベント間隔とイベント時降雨量は指数分布に従うのか？」なんて文章書きましたが、その中では、ポアソン過程と言っておきながら登場するのは指数分布ばかりでした。その説明が無いのは、なんとも気持ち悪いのではないかな？と思ったので、ずばり、ポアソン分布で表現されるポアソン過程は、表現の仕方次第で指数分布やガンマ分布でも表現されるということを紹介します（ガンマ分布による表現は後編で）。なお、本稿を書くのには、尾崎（1996）による教科書を参考にしました。

ある時刻  $t$  までに起こる事象の数を  $N(t)$  として、これを計数過程と呼びます。例えば、ある時間内の、店を訪れる客の数、かかってくる電話の本数、機械の故障の回数・・・など、いくらでも計数過程を想像できそうです。

ポアソン過程は、その計数過程で、これがポアソン分布で表現できることから、その名がつけました。ポアソン過程は、次のような性質を持ちます。と言うか、こういう性質を持つ確率過程をポアソン過程と呼びます。

1. 時間ステップ間  $(t_n - t_{n-1})$  での増分  $(N(t_n) - N(t_{n-1}))$  は独立であり、 $t_n$  と  $t_n+h$  ( $h$  は任意) に対する確率変数の分布  $(N(t_n)$  と  $N(t_n+h)$ ) は等しい。
2. 微小時間  $h$  の間に事象が一回起こる確率は、 $h$  に比例する。つまり、
$$P\{N(h) = 1\} = \lambda h + o(h) \quad (\lambda \text{ は比例定数、} o(h) \text{ は微小誤差})$$
3. 微小時間  $h$  の間に事象が二回以上起こる確率は、ほとんど無い。つまり、
$$P\{N(h) \geq 2\} = o(h)$$
4.  $N(0) = 0$

んんん、言葉による説明の方が、よく意味が分かんないですね。これから数式を追っていった方が分かりやすいと思うので、今、分からなくても無視して進んで下さい。

ずばり、時刻  $t$  までに  $k$  回の事象が起きる確率  $P_k(t)$  を考えます。すると、 $i$  を任意として、上記条件 1～4 より ( $h$  が微小時間なのを再確認して下さい)、

$$\begin{aligned} P_k(h) &= P\{N(h) = k | N(0) = 0\} = P\{N(t+h) - N(t) = k | N(t) = i\} \\ &= \begin{cases} 1 - \lambda h + o(h) & (k = 0) \\ \lambda h + o(h) & (k = 1) \\ o(h) & (k > 1) \end{cases} \quad (1) \end{aligned}$$

となります。余裕ある人には有効な情報ですが、 $1 - \lambda h + o(h)$  は  $e^{-\lambda h}$ 、だから当然、

$\lambda h + o(h)$  は  $1 - e^{-\lambda h}$  のテイラー展開の形になっているのを考えれば、「 $h$  が大きいと  $k = 0$  となる望みはない。」っていうのと「 $h$  が大きいと  $k = 1$  は絶対起きる。」っていうのが分かりやすいと思います。

で、 $P_k(t)$  の漸化式を作ります。  $k = 0$  となる場合を分けて考える必要があります。

$$\begin{aligned} P_0(t+h) &= P_0(t)P_0(h) \\ &= P_0(t)[1 - \lambda h + o(h)] \end{aligned} \quad (2)$$

$$\begin{aligned} P_k(t+h) &= P_{k-1}(t)P_1(h) + P_k(t)P_0(h) + \sum_{i=2}^k P_{k-i}(t)P_i(h) \\ &= P_{k-1}(t)[\lambda h + o(h)] + P_k(t)[1 - \lambda h + o(h)] + \sum_{i=2}^k P_{k-i}(t)o(h) \end{aligned} \quad (3)$$

もう、そんなに難しくありません。(2)、(3) 式は、 $h$  が微小時間であることを考えて、それぞれ次の微分方程式 (4)、(5) 式になります。ここで、 $o(h)$  は無視されました。

$$\frac{P_0(t+h) - P_0(t)}{h} = \frac{dP_0(t)}{dt} = -\lambda P_0(t) \quad (4)$$

$$\frac{P_k(t+h) - P_k(t)}{h} = \frac{dP_k(t)}{dt} = \lambda P_{k-1}(t) - \lambda P_k(t) \quad (5)$$

境界条件として  $P_0(0) = 1$ 、 $P_k(0) = 0$  を利用して、 $k = 0, 1, 2, 3, \dots$  という風に実験的に解いて行きます。本来なら、数学的帰納法が必要ですが、ま、ここでは無しということ。

すると、こんな式が得られます。

$$P_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t} \quad (6)$$

はい、これは、ずばり、パラメータ  $\lambda t$  のポアソン分布ですね。もう一度、何を求めたのか確認しましょう。(6) 式の表現するものは、「時刻  $t$  までに  $k$  回の事象が起きる確率  $P_k(t)$ 」です。パラメータ  $\lambda t$  のポアソン分布ということは、平均値と分散が  $\lambda t$  になりますから、「時刻  $t$  で起きた事象回数の平均は  $\lambda t$  回」と言えます。すると、 $\lambda$  は「時間当りに起きる事象の平均回数」を表現しているということになります。

さて、もうちょっと頑張りましょう。ポアソン過程に従う計数過程  $N(t)$  で、事象が起こる時間間隔を  $\tau$  とすると、その確率分布関数  $F(\tau)$  は「 $N(t+\tau) - N(t) = 0$  とならない」確率 (= 時間  $\tau$  の間に、必ず何か起きる確率) で表現されます。すると、(1) 式、

$$\begin{aligned} F(\tau) &= 1 - P\{N(t+\tau) - N(t) = 0\} = 1 - P\{N(\tau) = 0\} \\ &= 1 - P_0(\tau) = 1 - e^{-\lambda \tau} \end{aligned} \quad (7)$$

と書けます。すると、 $\tau$  の確率密度関数  $f(\tau)$  は、

$$f(\tau) = \frac{dF}{d\tau} = \lambda e^{-\lambda\tau} \quad (8)$$

となります。はい、これは指数分布ですね。くどいですが確認しましょう。ポアソン過程に従う計数過程を考えたとき、その事象の時間間隔  $\tau$  はパラメータを  $\lambda$  (=平均時間間隔) とする指数分布で表現できます。

#### 引用・参考文献

尾崎 (1996) 確率モデル入門. 朝倉書店.

□