

混合ポアソン過程の高波資料への適用

名古屋工業大学大学院

○河井 敦

名古屋工業大学大学院 正会員 北野 利一

1. 研究の目的 極大値資料から確率波高を算定する際、単位時間における平均出現数が重要となる。毎年最大値資料とは異なり、高波の超過確率を時間量と関連づけるのに、単位時間あたりの平均出現数が不可欠な量となるためである。これまでの極値波浪解析法においては、単位時間あたりの平均出現数を算出し、年間出現数の確率的変動特性をポアソン分布で与えている。しかしながら、年間出現数の確率分布をポアソン分布とする前提を、疑問視した研究は皆無といえる。このことは、波高の母分布に対して標本との適合性が入念に検討されることと対照的である。

そこで、本研究では、①ポアソン分布の特性を活用し、高波出現数の標本とポアソン分布との適合性について2種類の検討法を示す。②出現数の分布として、ポアソン分布以外に負の二項分布も適用できる確率過程モデルを示し、その統計的特性を考察し、高波資料への適用を試みる。

2. 研究の内容 解析には、秋田港の1991～2000年までの10年間における高波資料を用いる。極値理論より、閾値4.5mに対する超過波高について、極大波高の漸近分布である一般化パレート分布を適用した解析が可能である。

ポアソン分布は平均が比較的大きければ正規分布に近似できること、ポアソン分布の平均と分散は等しいこと、その標本平均と標本分散は独立であること、の3点の統計的性質を利用し、標本とポアソン分布との適合性の検討を行う。 k を観測出現数、 $\hat{\lambda}$ 、 $\hat{\sigma}^2$ をそれぞれ観測年数 N 年間の出現数の標本平均、標本不偏分散とすると、各年の観測出現数がポアソン分布に従うとき、標本不偏分散と標本平均の比 $\hat{\sigma}^2/\hat{\lambda}$ は、自由度で規格化したカイ自乗分布に従う。秋田港での高波資料について、これを利用した検定を行うと、平均と分散が等しいとする帰無仮説は、式(1)で示すように、有意水準5%で棄却される。

$$\chi^2(0.95, df=10-1)=16.9 < \chi^2_{\lambda} = 23.0 \quad (1)$$

したがって、ポアソン分布を高波出現数の母分布とすることは疑わしい。

二つ目の検討法として、高波の出現周期に着目する。年間出現数の標本が従う母分布を、ポアソン分布と仮定した場合、出現周期は、指数分布に従う。標本のquantileは、指数確率紙上ではほぼ直線上に配置されることが期待されるが、高波資料に関して、現実にそのような傾向を示すことは多くない。

そこで新たに、ポアソン分布の母数を出現強度とし、確率変量として扱い、出現強度 Λ の時間的な変動を考慮した確率過程、混合ポアソン過程(Reiss・Thomas, 2001)を検討する。図-1は、混合ポアソン過程を模式的に表したものである。この図では1年を2つに区分し、半期毎に出現強度が変動し、半期の区間内では、高波の出現は単純ポアソン過程と考え、各出現強度に応じて実現値が確率的に与えられることを示している。

ここでは、着目する期間(例えば、1年間)に疎密の周期に応じた副期間の数を(1年間で夏と冬の2期に分割する場合は、 $M=2$)とし、各副期間における出現強度の変動特性をガンマ分布 g (母数 α 、 β は発生数の平均および疎密により定まる)で与えれば、副期間における出現数の確率分布 $P'_{sub}(k)$ は、式(2)のように負の二項分布として得られる。

$$\begin{aligned} P'_{sub}(k) &= \int_0^{\infty} g(\lambda) P(k; \lambda/M) d\lambda \\ &= \frac{\Gamma(k+\alpha)}{k! \Gamma(\alpha)} \left(\frac{\beta M}{1+\beta M} \right)^{\alpha} \left(\frac{1}{1+\beta M} \right)^k \quad (2) \end{aligned}$$

負の二項分布の再生性により α を αM に置き換えると、着目する期間の出現数の確率分布 $P'(k)$ は、

$$P'(k) = \frac{\Gamma(k+\alpha M)}{k! \Gamma(\alpha M)} \left(1 + \frac{1}{\beta M} \right)^{-\alpha M} \left(\frac{1}{1+\beta M} \right)^k \quad (3)$$

と得られる。出現強度の分散が小さくなる($\beta \rightarrow \infty$)につれ、出現数の分散・平均比は1に近づく。それと同時に、出現数の分布 $P'(k)$ は、ポアソン分布に近似

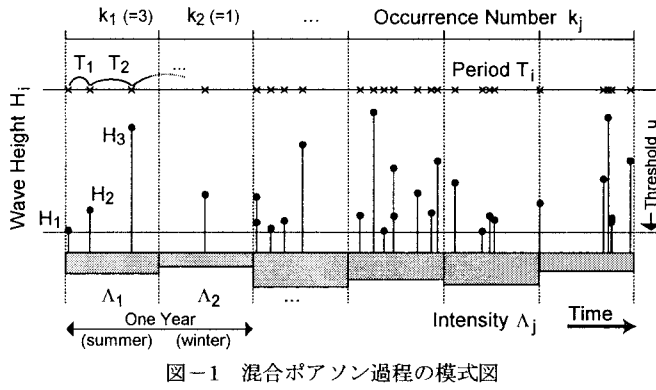


図-1 混合ポアソン過程の模式図

され、混合ポアソン過程が単純ポアソン過程を含むことを確認できる。

出現周期を考える場合には、出現強度の表現を出現周期あたりの出現数に改める必要がある。したがって、ここで扱う出現強度の分布 $g'(\lambda)$ は、

$$g'(\lambda) \propto g(\lambda) \cdot \lambda = g(\lambda; \alpha + 1, \beta) \quad (4)$$

となり、混合ポアソン分布での出現周期の確率密度は、

$$\begin{aligned} p'(\tau) &= \int_0^\infty g'(\lambda) p(\tau; \lambda) d\lambda \\ &= \frac{1}{\sigma} \left(1 + \xi \frac{\tau}{\sigma} \right)^{-(1+1/\xi)} \end{aligned} \quad (5)$$

と得られる。ここで、 $\sigma = \beta/(\alpha + 1)$; $\xi = 1/(\alpha + 1)$ となる。また、式 (5) はパレート分布として知られ、指数確率紙上で下に凸となる曲線を描く ($\xi = 0$ の場合は指数分布)。図-2 における点線は $\xi = 0.3, 0.6, 0.9$ の場合のパレート分布を描いたものである。出現周期の平均は、式 (6) のように得られ、平均出現数の逆数に一致する。

$$T = \frac{\sigma}{1 - \xi} = \frac{\beta}{\alpha} \quad (6)$$

秋田港での高波資料に対し、指数分布を母分布とした推定値 $\hat{T}_0$ 、パレート分布を母分布とした推定値 $\hat{T}'$ は、それぞれ

$$\begin{aligned} \hat{T}_0 &= 42.4 \pm 9.0 \quad [\text{day}] \\ \hat{T}' &= 49.0 \pm 31.1 \quad [\text{day}] \end{aligned} \quad (7)$$

と得られる。 $\hat{T}_0$ に比して、 $\hat{T}'$ は、そのとりうる範囲が広いため、パレート分布で推定するよりも、指数分布で推定する方が誤差は少ないと考えられる。しかし、この判断は、パレート分布を母分布とする立

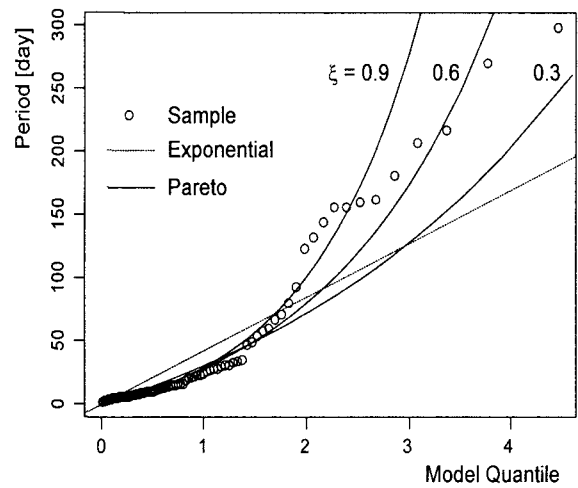


図-2 指数確率紙上の出現周期

場から見れば、第2種の誤りであるといえる。

3. 結論 従来の波浪極値解析における高波出現のモデルは、単純ポアソン過程モデルである。しかしながら、出現数の疎密の周期的な傾向が見られる場合には、出現強度の変化を考慮しなければならない。そこで、単純ポアソン過程を帰無仮説とする検定を提示し、棄却される場合に検討すべき確率過程モデルとして、混合ポアソン過程を提案した。

混合ポアソン過程では、年間出現数は負の二項分布に従い、その母数は各区間の出現強度の変動特性（期待値および分散）と年間の分割区間数 M に支配されることが示された。また、出現周期の確率分布についてはパレート分布となり、その母数の1つである平均出現周期は、負の二項分布から得られる平均出現数の逆数に一致することが分かった。

極値の漸近理論では、波高の分布 $f(x)$ は、式 (8) で表されるように、出現数に関する項 $1 - F(u)$ と閾値 u を越える波高に関する項 $h(x - u)$ に分離される。

$$f(x) = h(x - u) \{1 - F(u)\} \quad (8)$$

本研究では、そのうち出現数に関する項を対象に議論を進めた。今後は、点過程理論 (Coles, 2001) を用い、高波の出現数と波高を連動させて扱うことが可能な確率過程モデルの、検討を行う必要がある。

参考文献

- Coles, S. (2001): An Introduction to Statistical Modeling of Extreme Values, Springer, 208p.
Reiss, R. -D. and M. Thomas (2001): Statistical Analysis of Extreme Values, Birkhäuser, 443p.