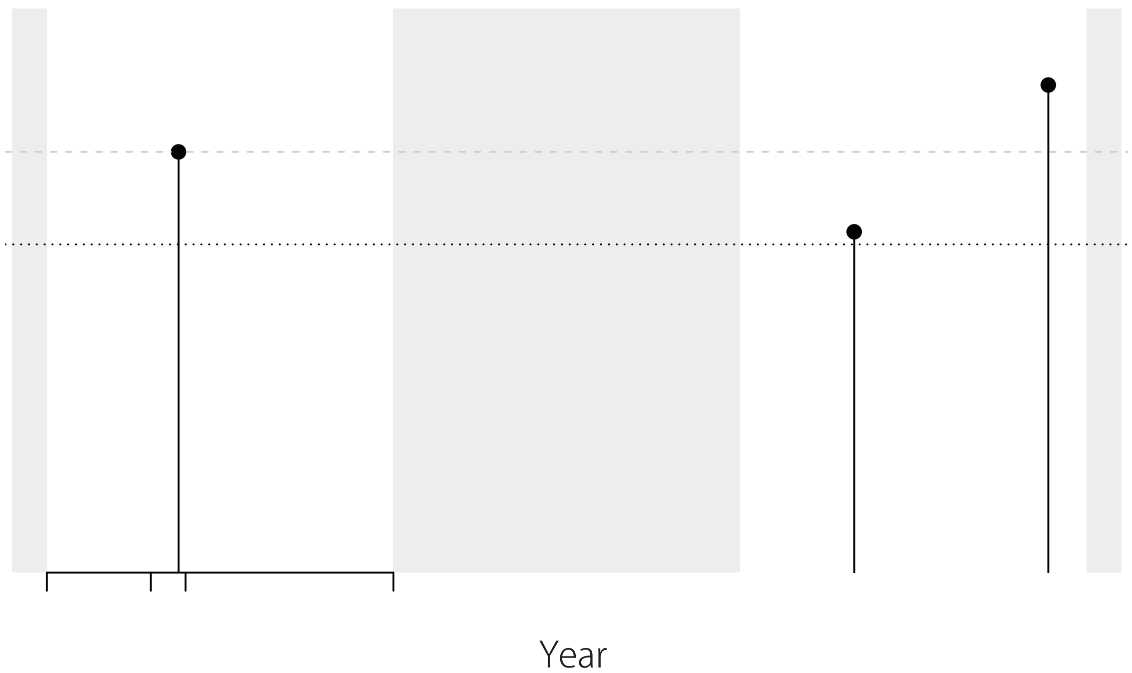
２０２０年３月１８日（水）12：30～17：30／日本学術会議講堂

日本学術会議公開シンポジウム
第９回防災学術連携シンポジウム「低頻度巨大災害を考える」

低頻度に伴う不確実性について ― 数理統計の視点から

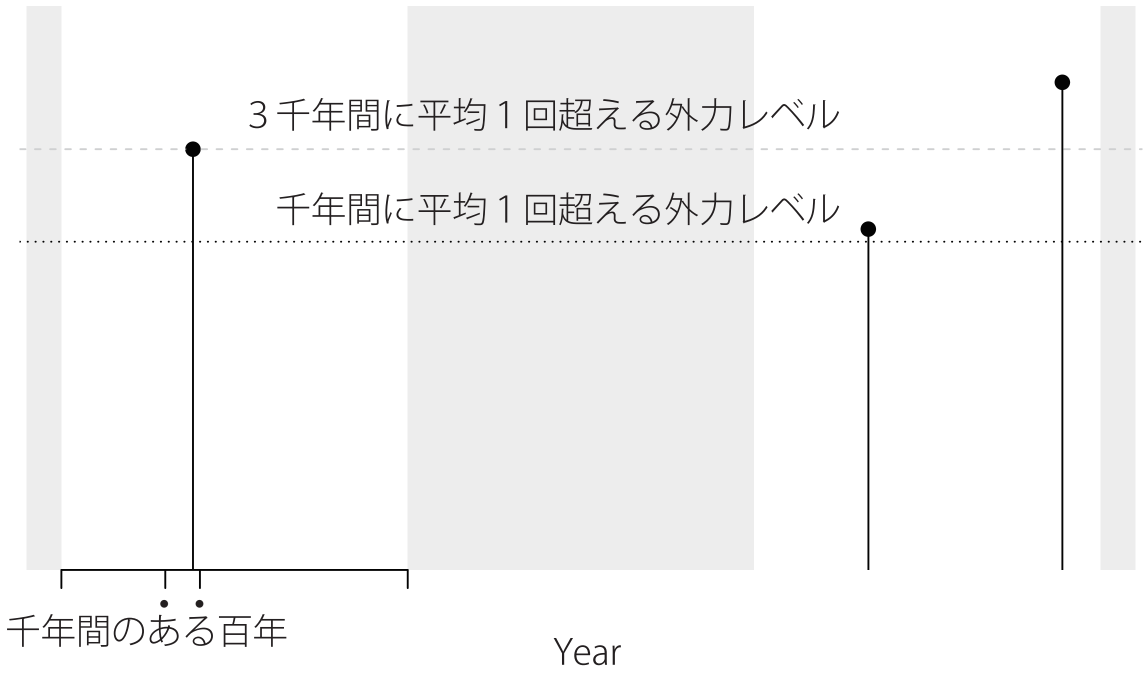
横幹連合　北 野 利 一
kitano@nitech.ac.jp
名古屋工業大学／統計数理研究所 客員

1

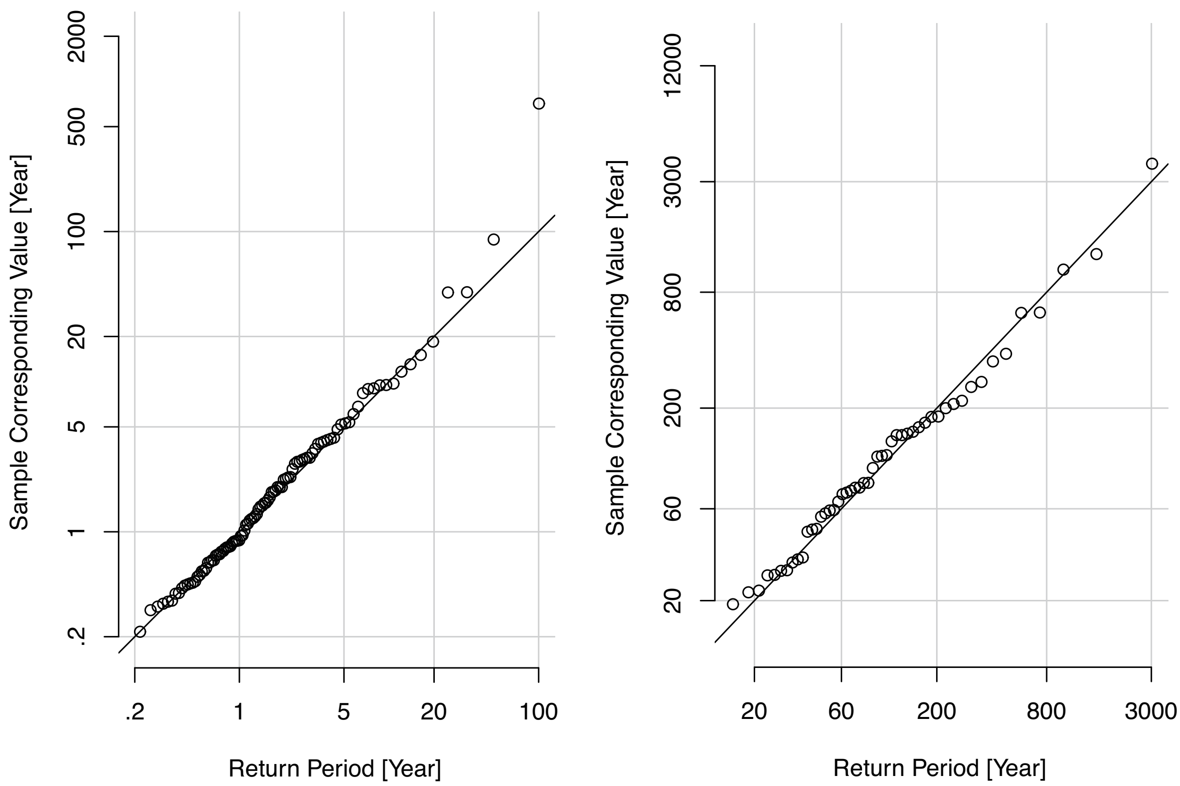


（意図的に，タテ軸そのものとヨコ軸の目盛の数字を付けていません）

2



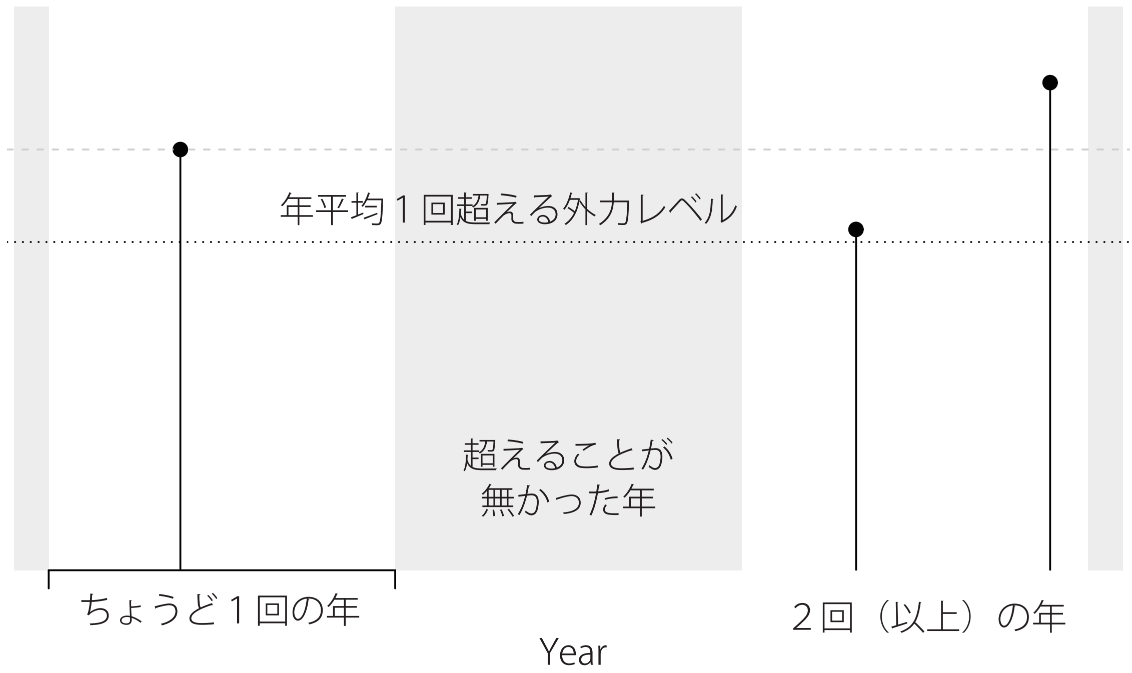
3



a) 100 年間の年最大値　　b) 60 年最大値（サンプルサイズ : 50）

外力に対応する再現期間（タテ軸）と期待される再現期間（ヨコ軸）
（外力の極値を再現期間の値に変換して得られる QQ プロット）

4



5

本日のメニュー

１）統計的な問題としての難しさ

（長い目で見て生じることが，わずかな観測期間内で生じたり，その逆に，
観測期間が，たまたま空白域に入ることもある．ヤッカイな問題である）

２）背景にある理論：　極値理論

（というより，ポアソン過程）

Keyword:　再現期間，　生起率，　３の法則と１／３の法則，
期間最大値モデル，閾値超過モデル，点過程モデル

３）推定に伴う誤差，外挿の限界

6

ポアソン分布から極値分布へ

1）ポアソン分布（1年間に平均 λ_1 回生起するイベントが，実際に k 回生起する確率）

$$f_1(k) = \frac{\lambda_1^k \exp(-\lambda_1)}{k!}$$

2）単純極値分布（再現期間 x 年を超えない外力の累積確率）

$$F_1(x) = \exp\left(-\frac{1}{x}\right) \quad \left(= \frac{\lambda_1^k \exp(-\lambda_1)}{k!}\bigg|_{k=0,\ \lambda_1=\frac{1}{x}} = e^{-\lambda_1}\bigg|_{\lambda_1=\frac{1}{x}} \right)$$

3）外力（例えば，降水量）y を超える生起率（1年あたりの生起数の期待値）

$$\lambda_1 = \frac{1}{x} = \left(1 + \xi\,\frac{y - \mu_1}{\sigma_1}\right)^{-1/\xi}$$

$$= \exp\left(-\frac{y - \mu_1}{\sigma_1}\right) \qquad (\xi = 0)$$

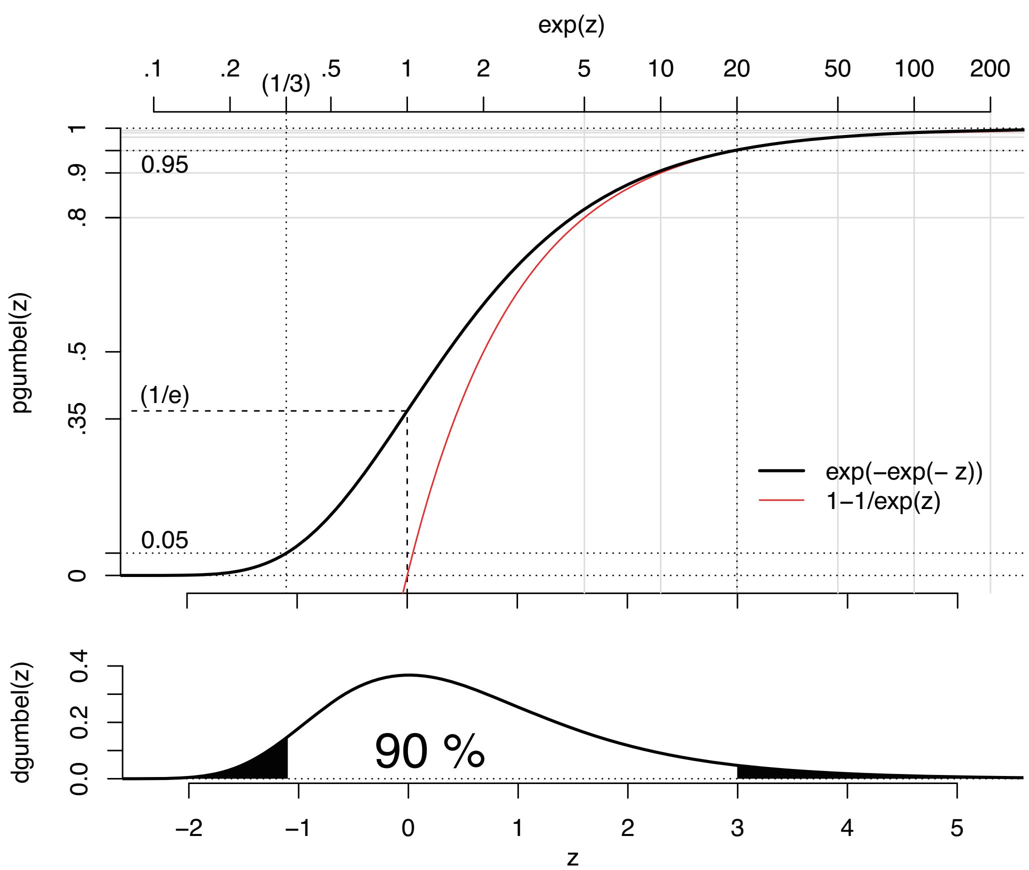
パラメータ μ_1, σ_1, ξ を現象固有の説明変数でモデル化も可能

注：この関係は，生起率の比例性を用いて，関数方程式を解くことにより導出（de Haan & Ferreira）

4）標準グンベル分布 $\quad F(z) = e^{-e^{-z}}$

（ガンベルと発音されることもしばしばあることに注意）

7



標準グンベル分布

8

標準グンベル分布 $F(z) = e^{-e^{-z}}$ に対する**９０%の範囲に隠されたヒミツ**

$$F(3) - F(-1) \approx F(3) - F(-\log 3) \approx e^{-e^{-3}} - e^{-3}$$
$$\approx \left(1 - \frac{1}{20}\right) - \frac{1}{20} = 1 - \frac{2}{20} = 0.9$$

★ ３の法則（Rule of three）

イベントが全く生じていない現実を見て，そもそも起こりえないと考えるのではなく，本来，平均 $\lambda = 3$ も生じることが，現実には，たまたま１度も生じていないだけと考える．

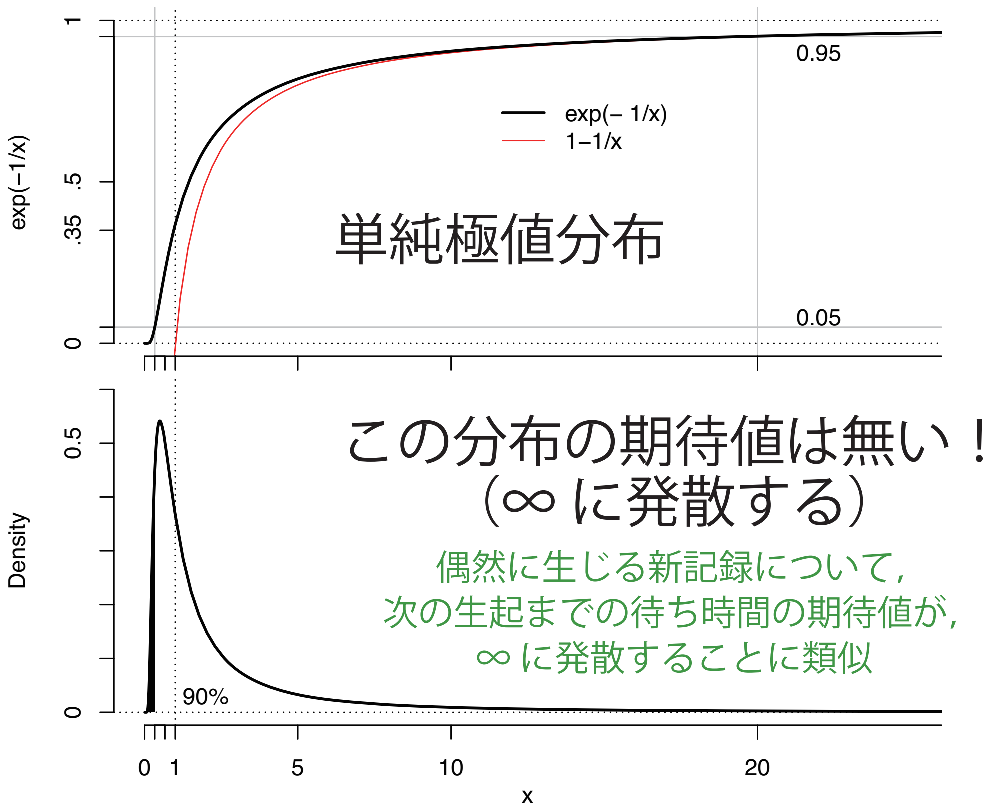
$$\left.\frac{\lambda^k \exp(-\lambda)}{k!}\right|_{\lambda=3,\, k=0} = e^{-3} \approx \frac{1}{20} \quad \leftrightarrow \quad \log 20 \approx 3$$

★ １／３の法則（Rule of thirds）

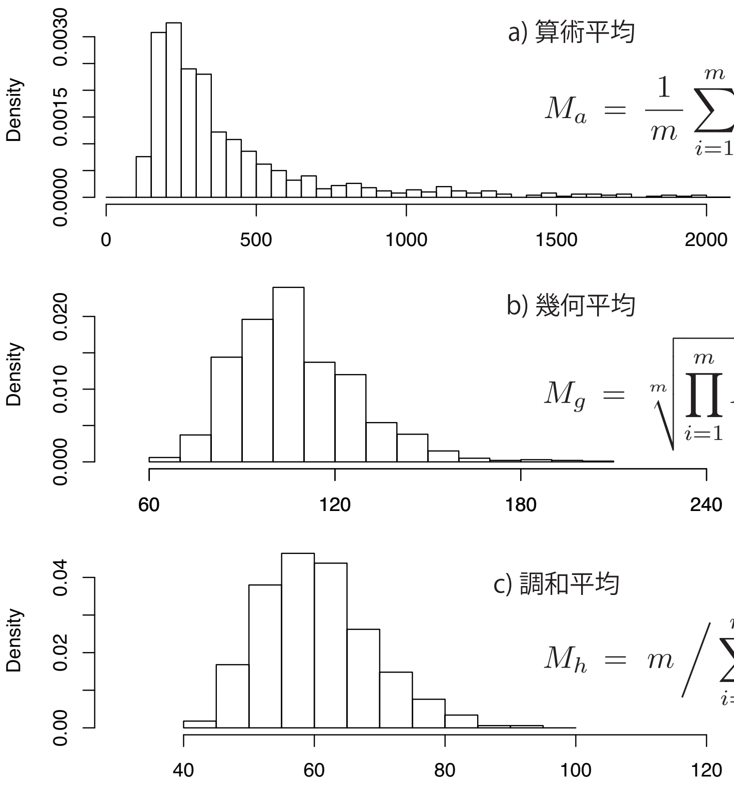
平均１回生じるイベントが生じない確率は，およそ 1/3 であり，むしろ，１回以上生じる確率は，残りの約 2/3 となり，生じない確率の約２倍である．

$$\left.\frac{\lambda^k \exp(-\lambda)}{k!}\right|_{\lambda=1,\, k=0} = e^{-1} \approx \frac{1}{3} \quad \leftrightarrow \quad \log 3 \approx 1$$

$$\sum_{k=1}^{\infty} \left.\frac{\lambda^k \exp(-\lambda)}{k!}\right|_{\lambda=1} \approx \frac{2}{3}$$



期間最大値に対する単純極値分布の累積確率とその確率密度
（期間最大値を再現期間に読み替えて，その再現期間の累積確率分布とも言える）



a) 算術平均

$$M_a = \frac{1}{m}\sum_{i=1}^{m} X_i'$$

b) 幾何平均

$$M_g = \sqrt[m]{\prod_{i=1}^{m} X_i'} = \exp\left\{\frac{1}{m}\sum_{i=1}^{m}\log X_i'\right\}$$

c) 調和平均

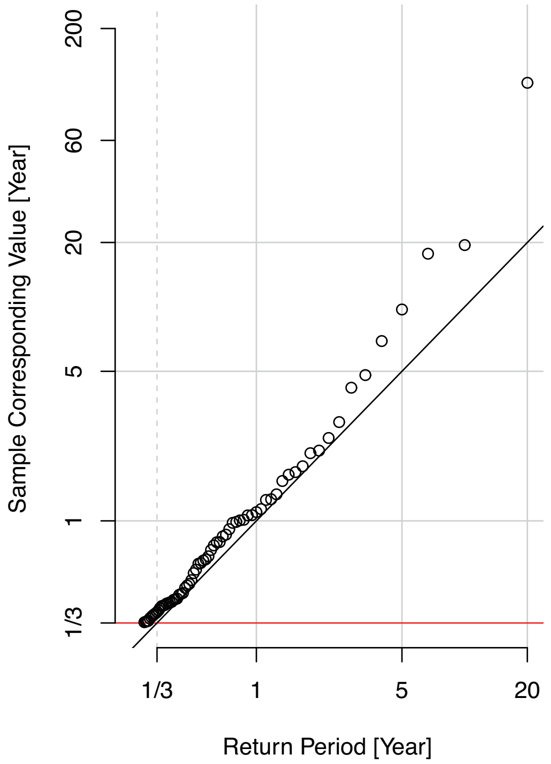
$$M_h = m \Big/ \sum_{i=1}^{m}\frac{1}{X_i'}$$

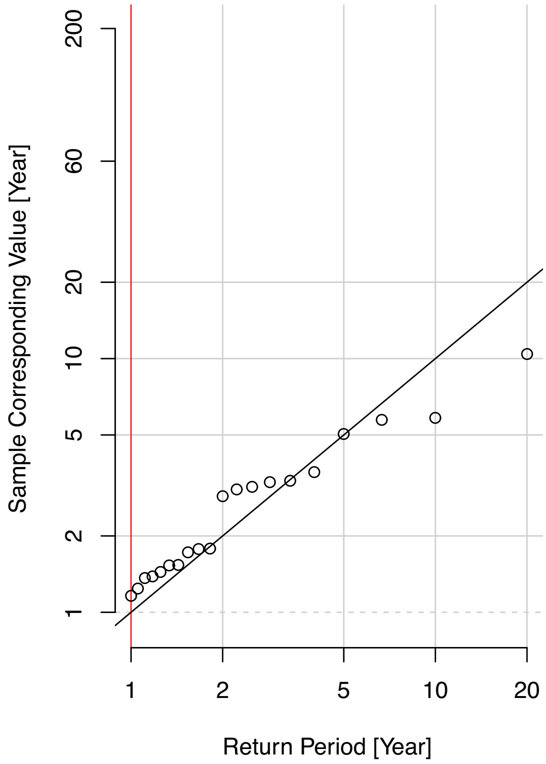
再現期間の期待値は
∞ に発散する！

単純極値分布からサンプリングされた３種類の標本平均

11

期間最大値のみならず，極大値も用いると，推定における情報が増える

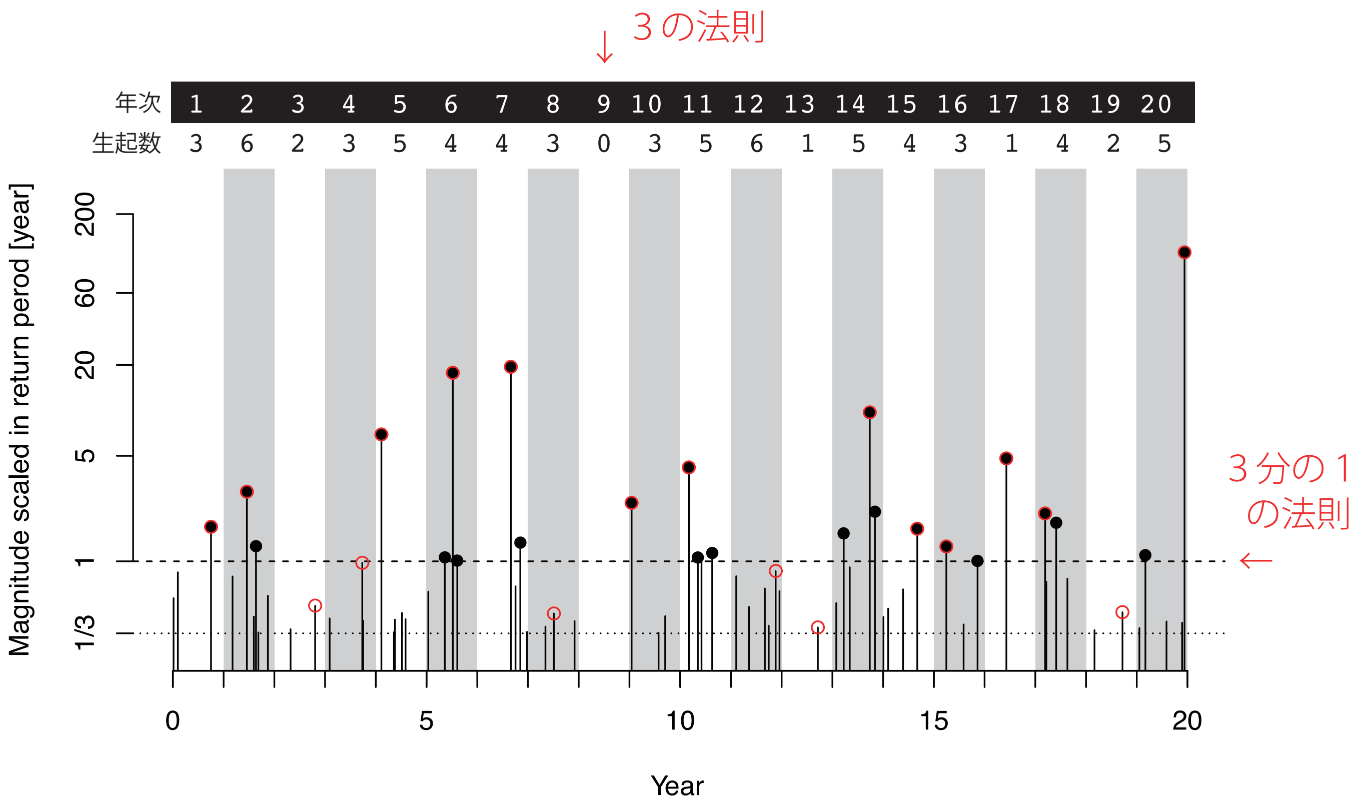




a) 閾値（生起率３）を超過する極大値　　b) 上位 20 番目までの極大値

外力に対応する再現期間と期待される再現期間
（a, b は異なる 20 年間に生起する極大値を抽出）

12



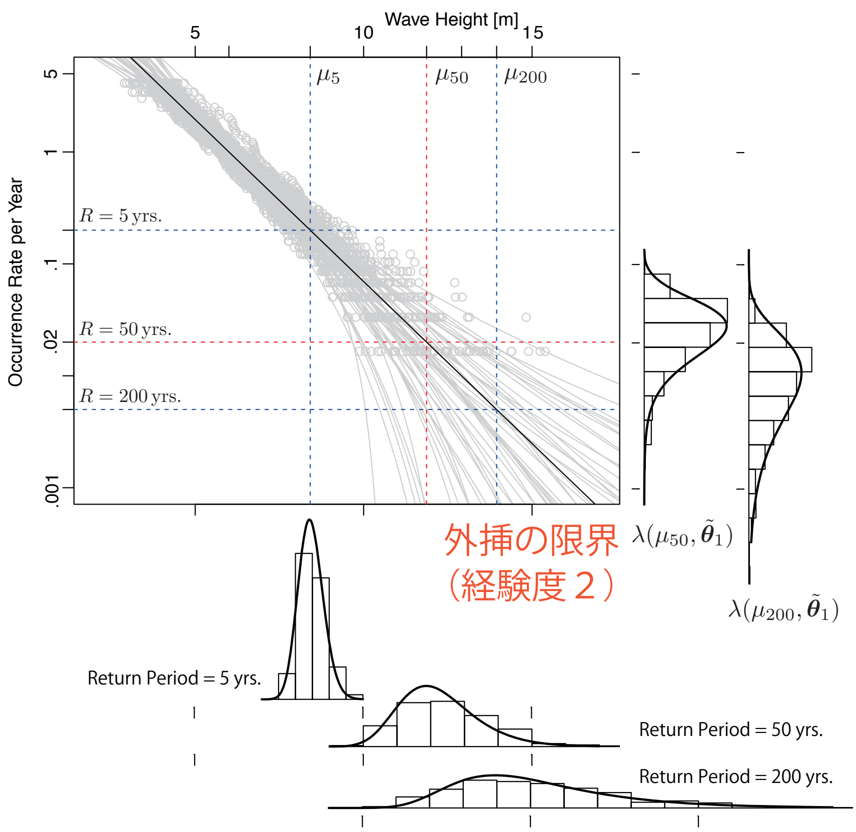
期間最大値モデル と 閾値超過モデル

いずれも，ポアソン過程（点過程モデル）に帰着

極値分布そのもののというよりも，ポアソン分布と生起率関数が重要である

13

順序統計量を用いた極値統計解析による推定（誤差）の解釈



所与のレベルを超える生起率の推定値の統計的変動と，
推定される再現レベルの統計的変動の関係

例えば，定常な３千年間の記録がある時，上位 30 番めの極値は，100 年確率外力値と推定できる．

もちろん，推定誤差を伴うことに注意しなければならない．

60 年間の観測記録を用いて，100 年確率外力値を外挿により推定した時に，経験度 2.1 を得た．これは仮想的に 210 年間の記録をもとに，上位 2.1 番めに相当する値を 100 年確率外力値と推定していることになる．

再現レベルの誤差分散：

$$\mathrm{V}(\hat{\mu}_n) = \frac{\sigma_n^2}{K}$$

cf.　標本平均の誤差分散：

$$\mathrm{V}(\overline{X}_n) = \frac{\sigma^2}{n}$$

14

本日のまとめ（メニューに加筆）

1）統計的な問題としての難しさ：神出鬼没性

（長い目で見て生じることが，わずかな観測期間内で生じたり，その逆に，
観測期間が，たまたま空白域に入ることもある．ヤッカイな問題である）

2）背景にある理論：　ポアソン過程
（極値理論は，その上屋にすぎない）

Keyword:　再現期間，　生起率，　３の法則と１／３の法則，
期間最大値モデル，閾値超過モデル，点過程モデル

3）推定に伴う誤差，外挿の限界

順序統計量を用いた考え方を紹介（2度あることは3回ある，と考えて備えよ）

4）２変量（多変量）極値は，別の機会に紹介

時間を空間でかせぐ方法，ただし，空間相関をキチンと扱うべきである

参考文献（さらに理解を深めるために，. . . ）

de Haan, L. and A, F. Ferreira: Extreme Value Theory - An Introduction,
Springer, 417p., 2006.

北野利一 : 伊勢湾台風級の高潮と確率潮位，
第 49 回 水工学に関する夏期研修会，土木学会，20p., 2013.

北野利一 : 恋する極値統計，統計数理研究所 共同研究リポート，433，
「極値理論の工学への応用（17）」，pp. 142-161, 2020.

Coles, S.: An Introduction to Statistical Modeling of Extreme Values,
Springer, 208p., 2001.

高橋倫也・志村隆彰 : 極値統計学，近代科学社，262p., 2016.