日本の降雨の*N*年間最大値の統計

～想定外の出来事を想定外としないための～

唐沢　好男

発表の概要

本発表は前発表の続編である。前発表では、我が国の強雨現象の経年変化を調べる目的で、気象庁が公開している気象データベース[2]から、各種降水量（日降水量・時間降水量・10分間降水量）の年間最大値を用いた統計解析を行った。日本全国45地点、最大100年（ただし、降水量によって地点数と年数は異なる）のデータを地点ごとの平均値で正規化した降水量（正規化降水量）について、95%信頼区間推定を行い、日本全体では、最大降水量が100年スケールで見て10%程度増加している傾向を読み取ることができた。ただし、最大降水量は年毎の変動が大きく、地点別に見ると100年という長期間であっても、その変化傾向が有意とは言えない程度のものであった。

そこで、本発表では、前発表で用いたと同じ正規化降水量を用い、全データが確率的に均一と仮定し、そのデータを用いて、10年、100年、さらには、1000年に一回の強雨量を求める。災害は忘れたころにやってくると言われるが、想定外の出来事を想定外としないための心構え（備え）に役立つ情報になる。この種の解析では、極値統計の理論（Gumbel分布）がデータ解析結果の解釈に力を発揮することになる。果たして、何が見えてくるであろうか。

発表の内容

１. 解析の道具

・最大値の統計

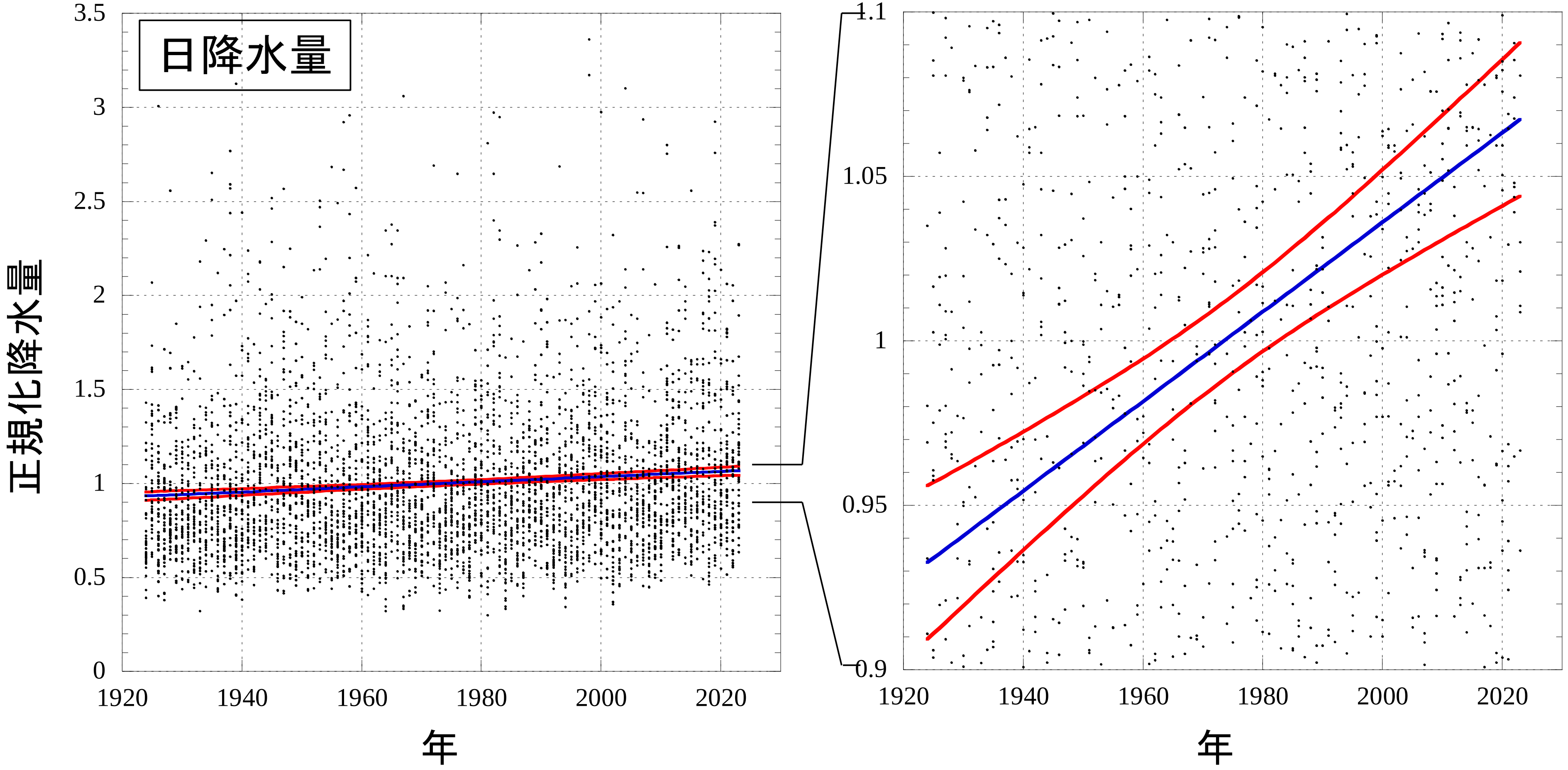
・極値統計

２. 気象データ（詳細説明略）

３. N年間での最大降雨

・多地点データで見る

・仮想１地点に拡大したデータで見る



長期的な傾向変化よりは、ばらつきのほうがはるかに大きいので、
データ全体を一つの定常確率過程から生まれたデータとして見てみよう

順序統計・最大値統計・極値分布

n 個の標本値（小さい順）

母集団

確率分布

$f(x)$

$F(x)$

独立試行

1
$$x_1^{(1)} \le x_2^{(1)} \le \cdots \le x_i^{(1)} \le \cdots \le x_n^{(1)}$$

2
$$x_1^{(2)} \le x_2^{(2)} \le \cdots \le x_i^{(2)} \le \cdots \le x_n^{(2)}$$

$$\vdots$$

順序統計

$f_{(i)}(x)$

$F_{(i)}(x)$

最大値統計

$f_{\max}(x)$

$F_{\max}(x)$

$n \to \infty$ $\Longrightarrow$ 極値分布（Gumbel 分布等）

順序 i 番目の分布と最大値の分布

母集団の確率分布： f : 確率密度関数、F : 累積分布関数

n 個のサンプル値の小さいほうから i 番目の値の確率分布（順序統計）

$$F_{(i)}(x) = \sum_{k=i}^{n} \binom{n}{k} F^{k}(x)\{1-F(x)\}^{n-k}$$

$$f_{(i)}(x) = n\binom{n-1}{i-1} F^{i-1}(x)\{1-F(x)\}^{n-i}\ f(x)$$

$$\binom{n}{k} \equiv \frac{n!}{k!(n-k)!} \qquad (0 \le k \le n)$$

最大値の分布（$i = n$ より）

$$F_{\max}(x) = F_{(n)}(x) = F^{n}(x)$$

$$f_{\max}(x) = f_{(n)}(x) = nF^{n-1}(x)\, f(x)$$

元（母集団）の分布形が
与えられれば
具体的な形に記述できる

最大値の分布の平均値

期待値

$$\left\langle x_{\max} \right\rangle = \int_{-\infty}^{\infty} x f_{\max}(x) dx$$

← 一般的には、数値積分

近似値

（最頻値に近い）

$$\left\langle x_{\max} \right\rangle \approx F_{\max}^{-1}\left(\frac{n}{n+1} \right)$$

← 累積分布の逆関数が閉形式で表されれば計算に便利な式

極値分布

最大値の累積確率分布

$$F_{\max}(x) = F_{(n)}(x) = F^n(x)$$

$n \to \infty$ の極限において、F の形により、3つの分布系（極値分布）に収斂する

x の値が大きい部分（分布の裾）が指数関数的なものは、Gumbel分布に収斂する。電波伝搬に現れる大部分の分布（正規分布、指数分布、ガンマ分布、ワイブル分布、対数正規分布など）の極値分布はGumbel分布になる

Gumbel分布

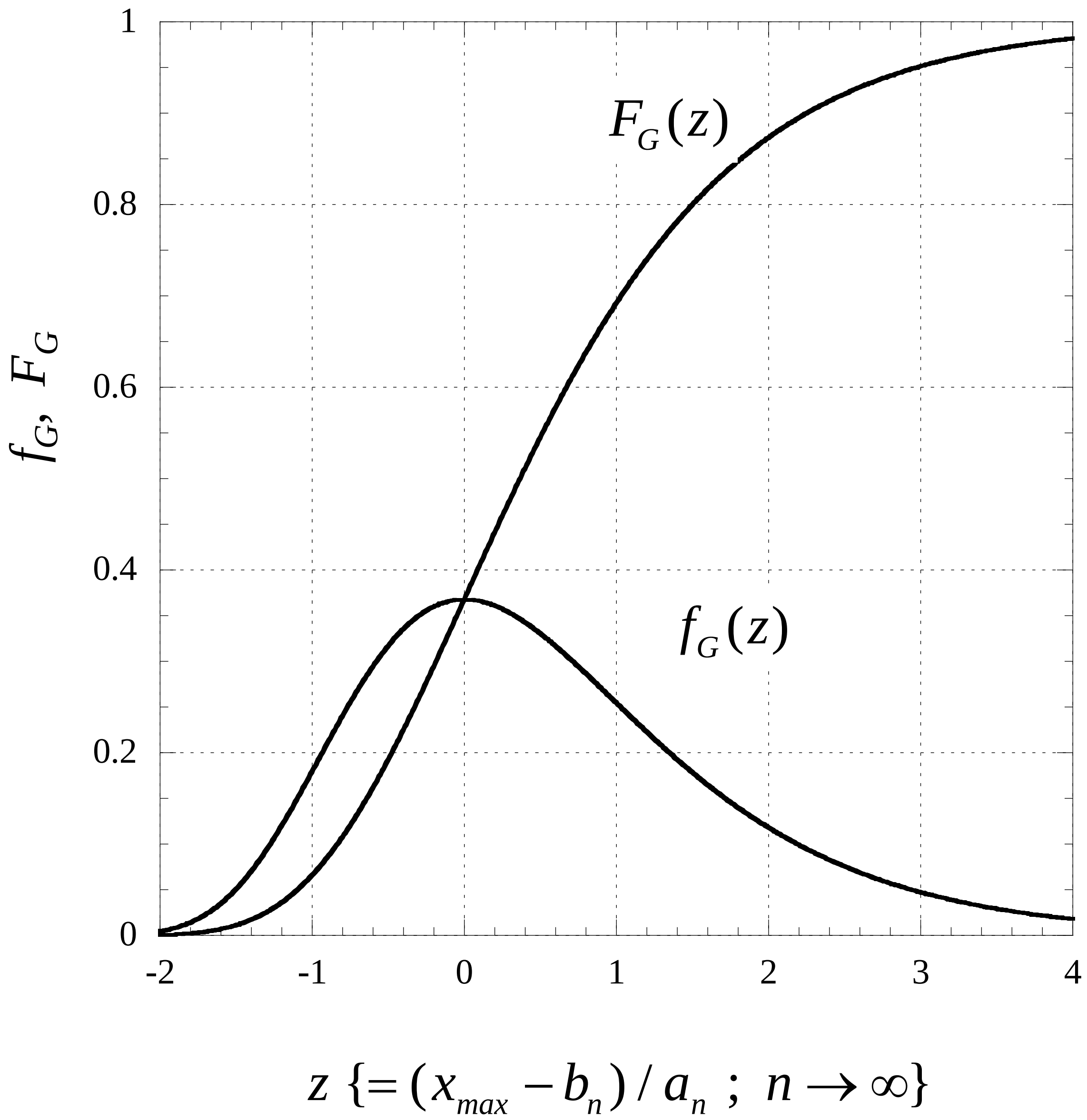
$$F_G(z) = \exp\{-\exp(-z)\}$$

$$f_G(z) = \exp\{-z - \exp(-z)\}$$

$$z = \frac{x_{\max} - b_n}{a_n}$$

分布のパラメータ a_n, b_n は吸引係数と呼ばれる

極値分布：Gumbel 分布の形



本研究の解析に用いた降水量データ（気象庁のHPより）

日本各地の降水量年間最大値

日降水量	45地点	100年間	（1924 - 2023）
時間降水量	45地点	80年間	（1944 - 2023）
10分間降水量	47地点	70年間	（1954 - 2023）

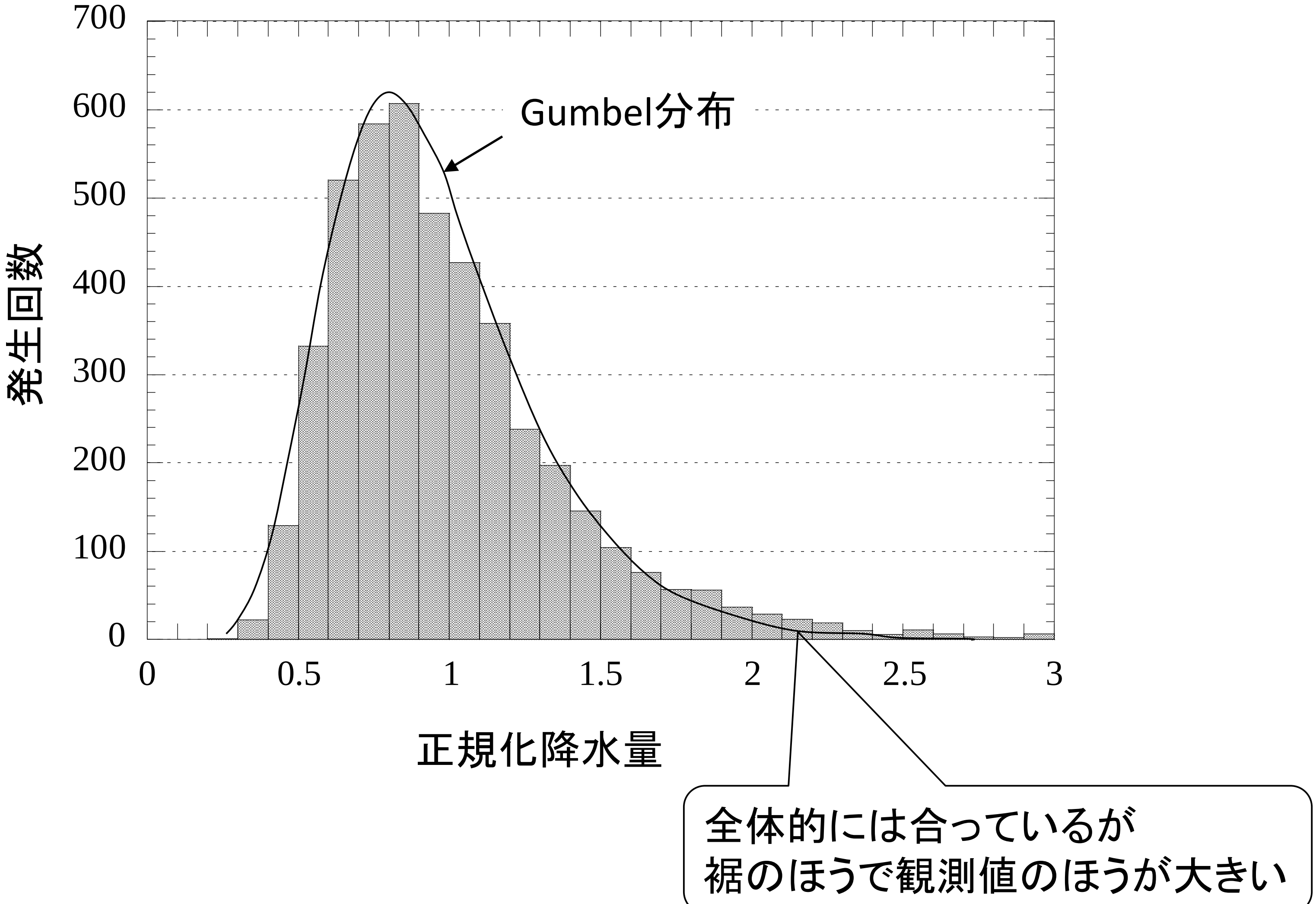
このデータをそれぞれの地点の平均値で正規化した正規化降水量を使う

年間の最大値であるので、

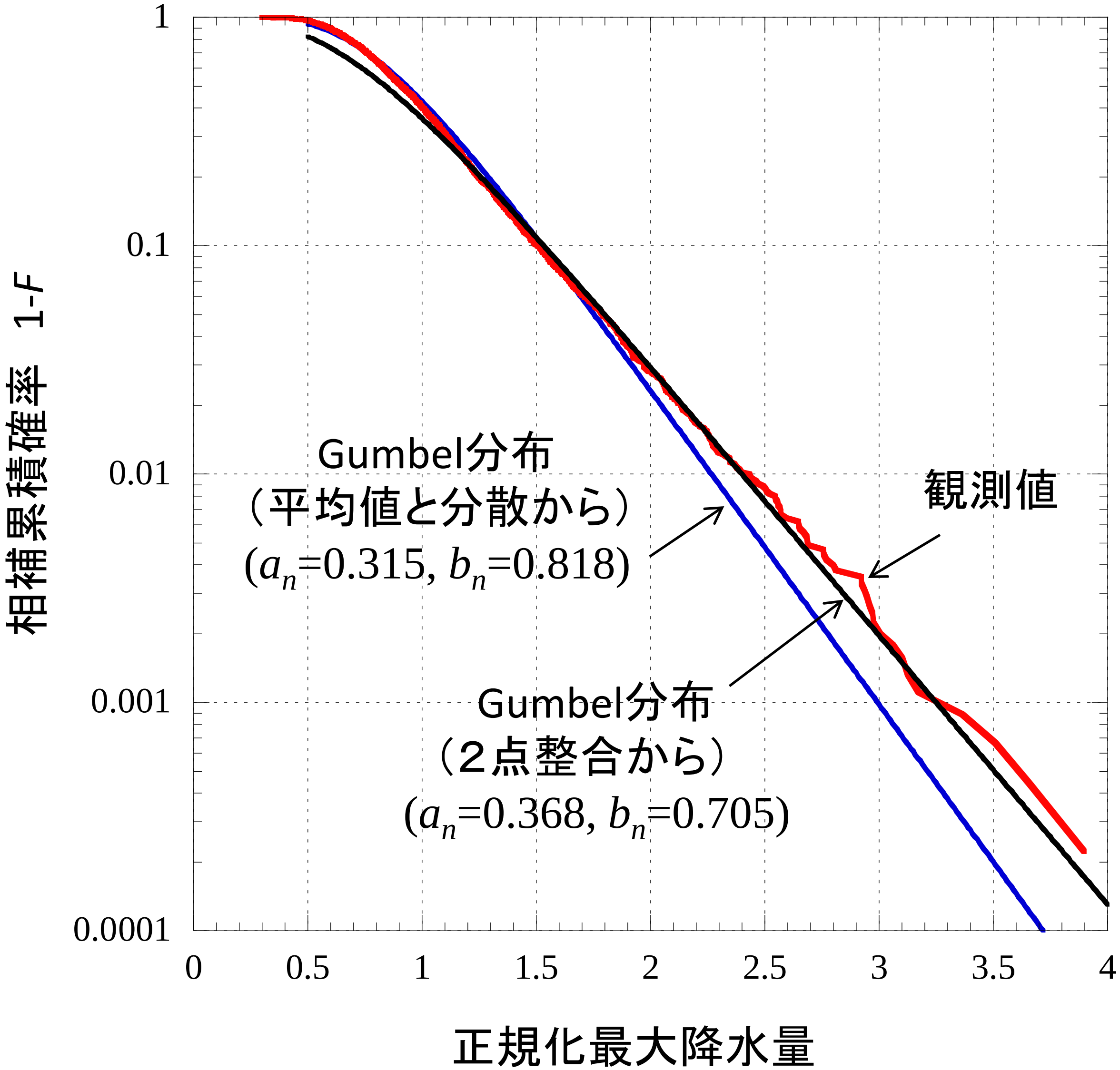
日降水量	$n = 365$
時間降水量	$n = 8{,}760$
10分間降水量	$n = 52{,}560$

十分大きい数といえるかどうか
特に、日降水量

正規化日降雨量年間最大値（4500個）のヒストグラムとGumbel分布でのフィッティング



正規化最大日降水量の相補累積分布とGumbel分布への近似



Gumbel 分布の平均値

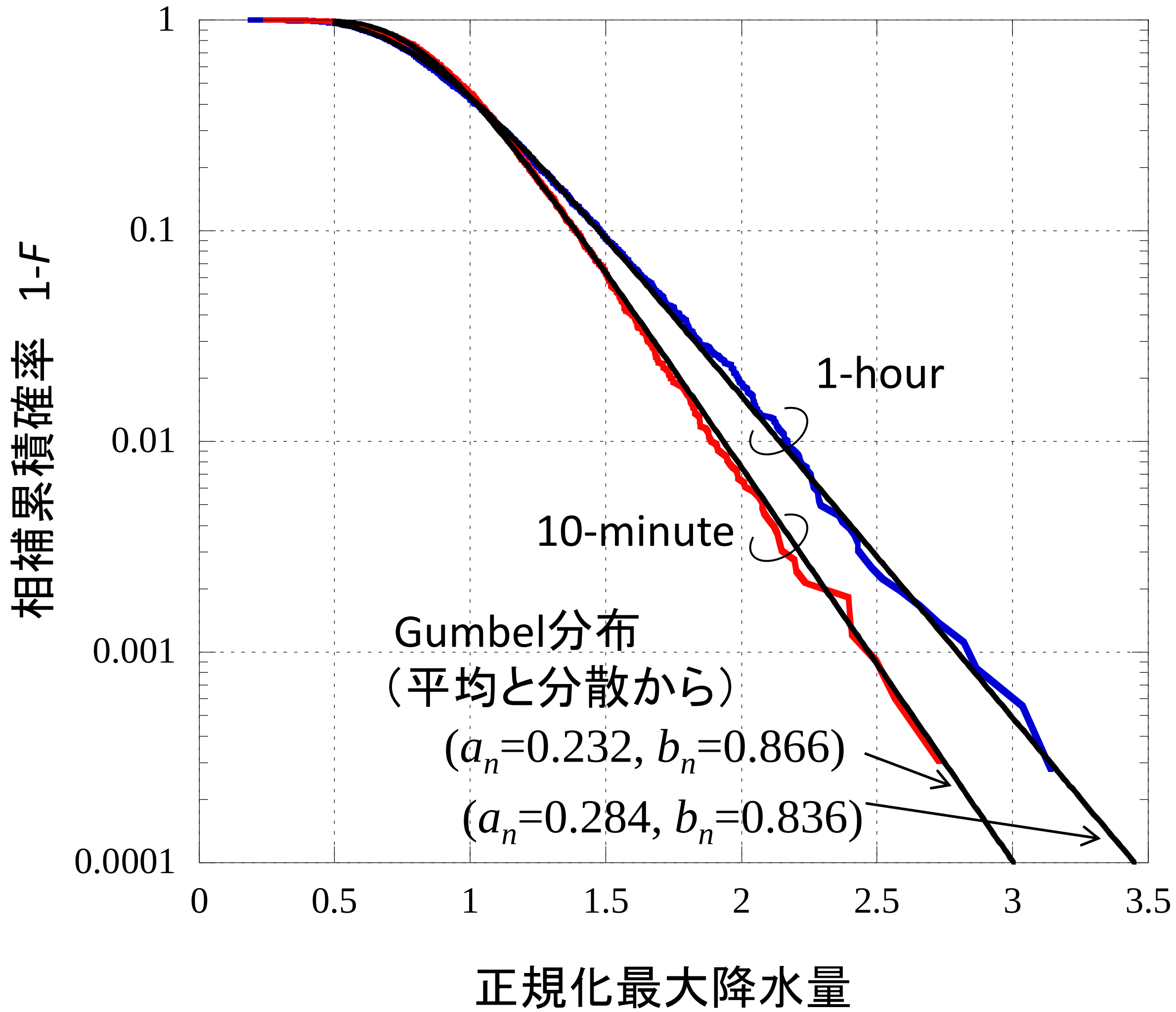
$$m_G = \langle x_{\max} \rangle = b_n + \gamma a_n$$

（γ: オイラーの定数 =0.5772・・・）

分散

$$V_G = \left\langle \left(x_{max} - m_G \right)^2 \right\rangle = \frac{\pi^2 a_n^2}{6}$$

正規化最大時間降水量と10分間降水量の相補累積分布とGumbel分布への近似



N年間に一度の最大値を求める

観測値

各地点のデータをN年ごとに分割してグループを作り、グループ毎の最大値を求めて平均値を出す
例えば、100年間のデータを持つ日降水量に対して、N=50の場合、地点ごとに2個、全部で90個の最大値が得られ、その平均値を求める。

極値分布からの推定

１）全データに対する吸引係数 $a_n,\ b_n$ を求める
（日降水量：２点整合法；　時間降水量と10分間降水量：平均値と分散から）

２）平均値を求める

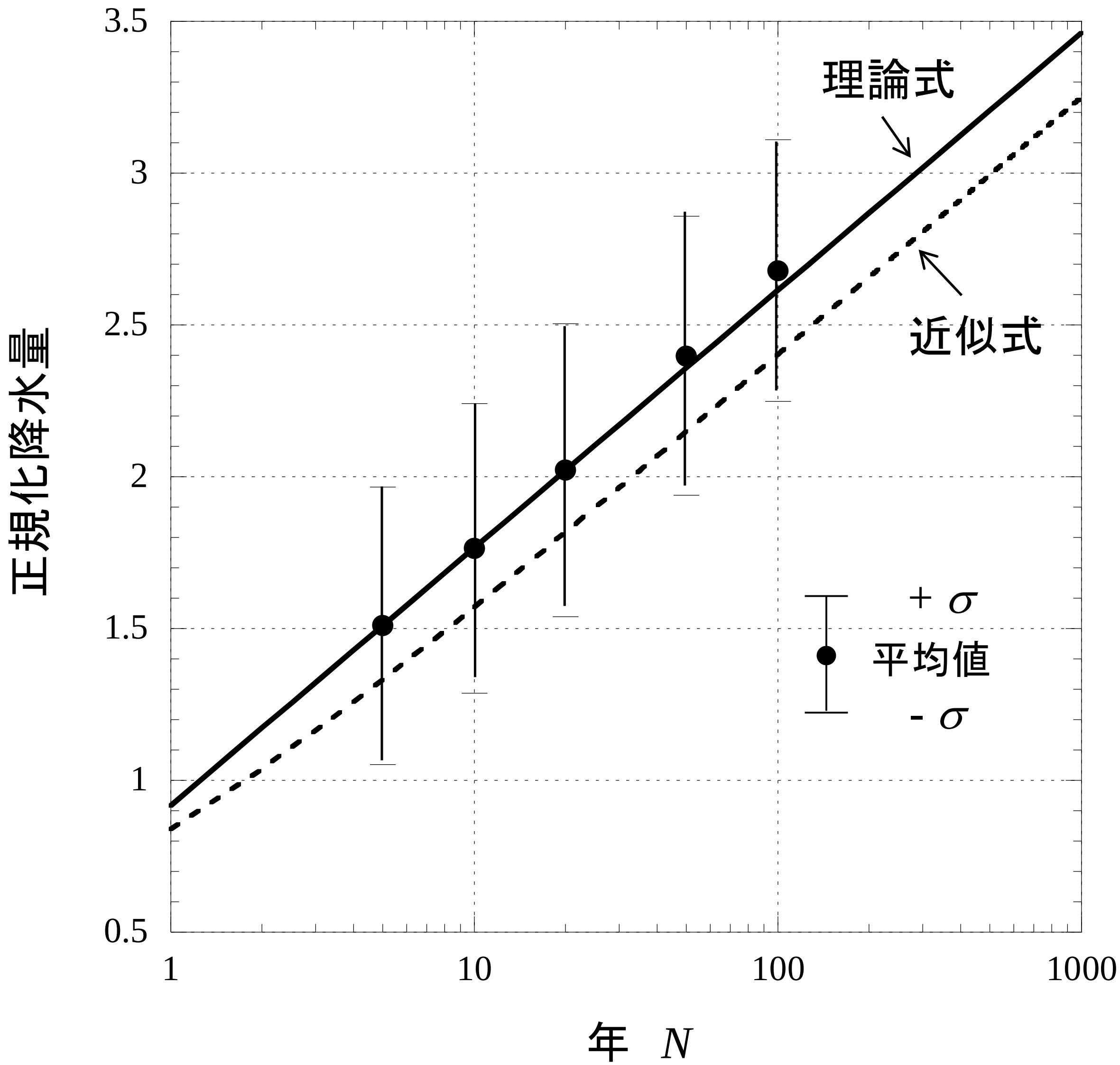
平均値（数値積分）

$$\langle x_{\max} \rangle = \int_{-\infty}^{\infty} x f_{\max}(x) dx \quad \Longrightarrow \quad x_N = \frac{N}{a_n} \int_0^{\infty} x \left\{ F_G\left(z(x)\right) \right\}^{N-1} f_G\left(z(x)\right) dx$$

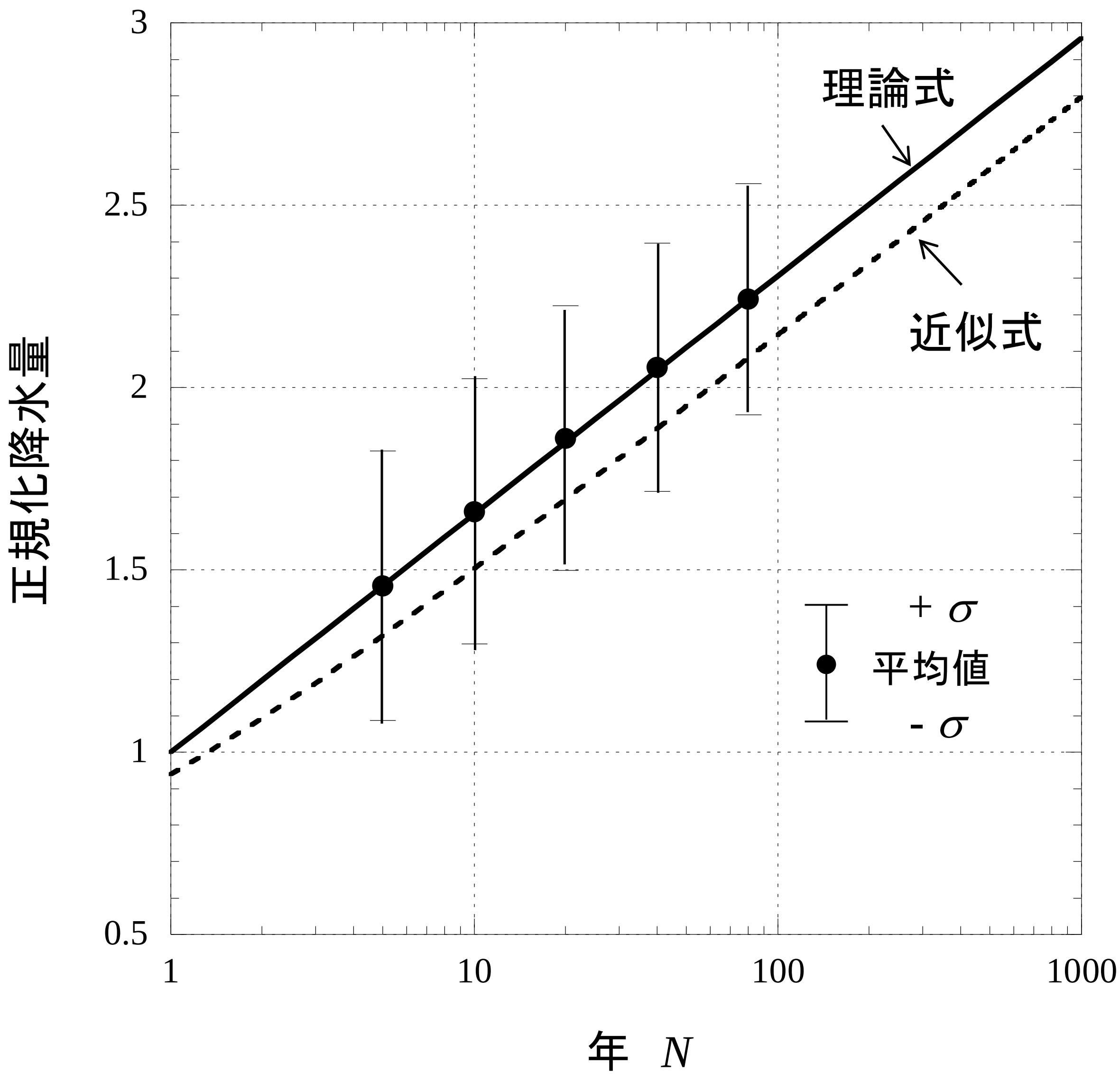
近似値（閉形式が得られる）

$$x_N \approx F_G^{-1}\left(\frac{N}{N+1}\right) = -a_n \log\left\{ -\log\left(\frac{N}{N+1}\right) \right\} + b_n$$

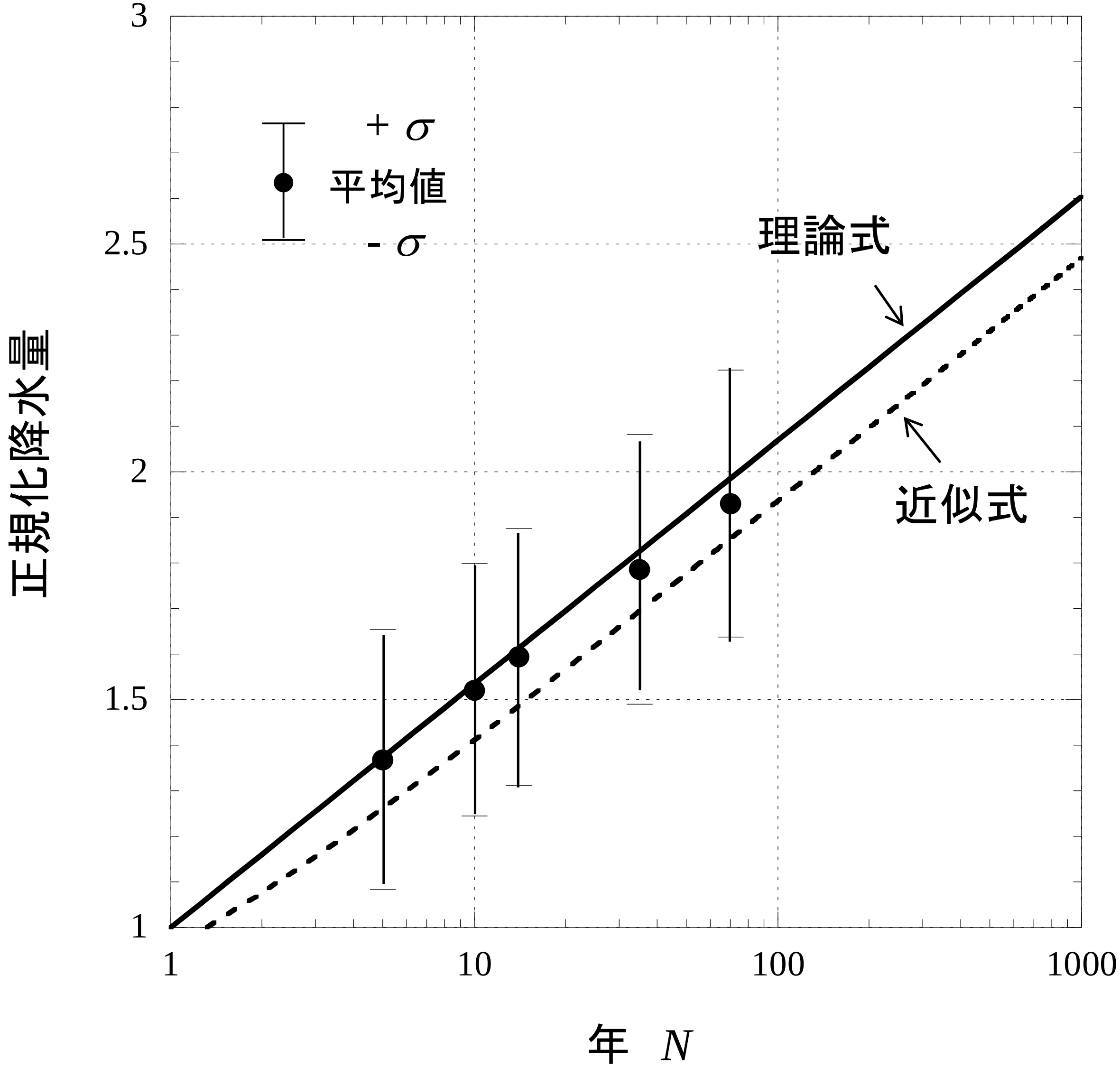
正規化日降水量のN年間最大値



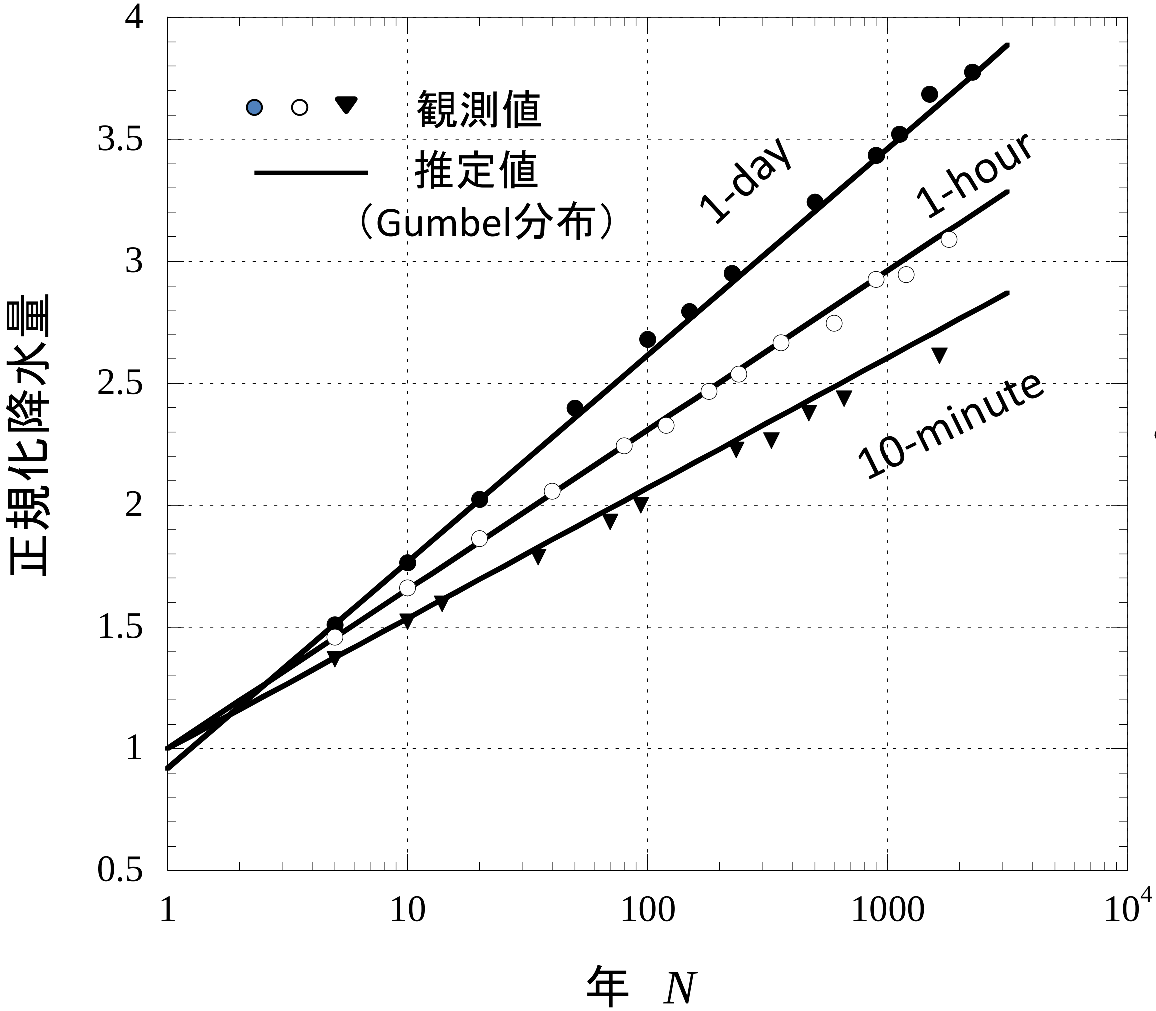
正規化時間降水量のN年間最大値



正規化10分間降水量の*N*年間最大値



多点正規化雨量をつなげて一点での仮想長期間データとして見てみよう



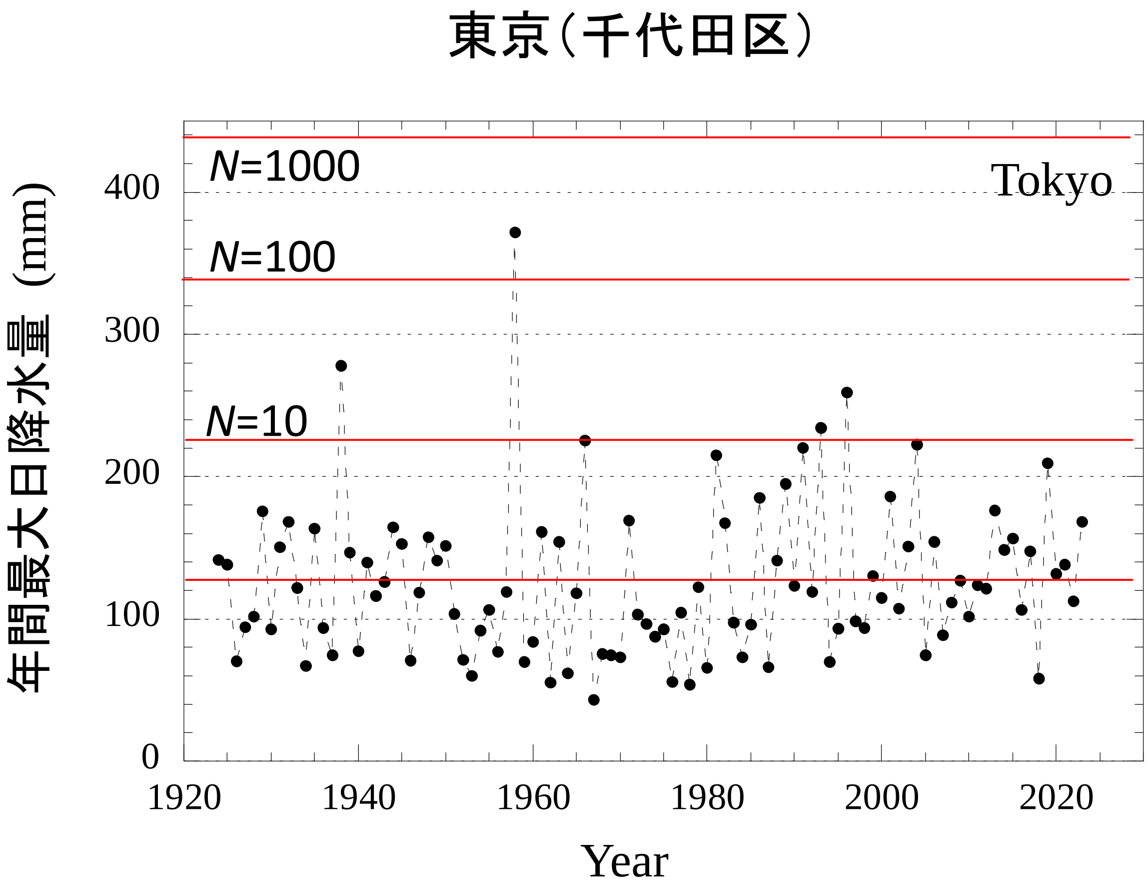
日降水量:　　4500年分
時間降水量:　3600年分
10分間降水量: 3290年分

Gumbel分布からの推定式

$$\left\langle R_{\max}(N)\right\rangle = \left(a\log_{10} N + 1\right)\left\langle R_{\max}(1)\right\rangle$$

$$a = \begin{cases} 0.83 & \text{(1-day rainfall)} \\ 0.66 & \text{(1-hour rainfall)} \\ 0.53 & \text{(10-minute rainfall)} \end{cases}$$

東京における年間最大日降水量の100年間の推移と*N*年間最大値の推定値



100年間平均値
125.7 mm

推定値

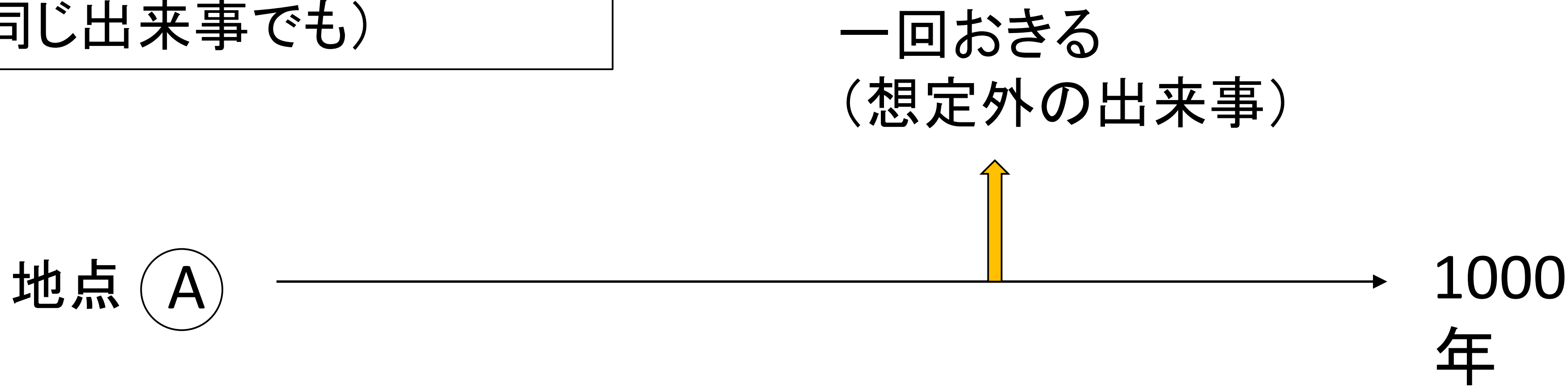
N=10　226 mm
N=100　339 mm
N=1000　440 mm

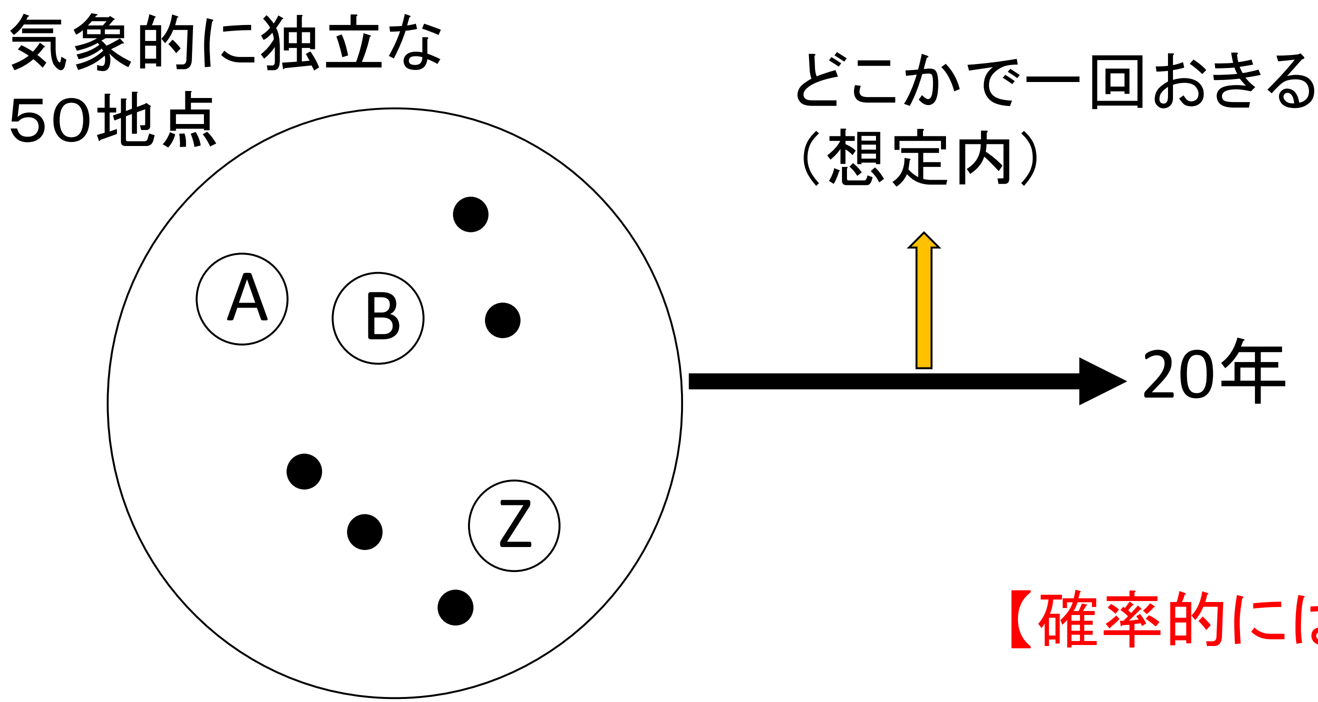
観測値

100年間最大
371.9 mm (1958)

2番目
287.3 mm (1938)

パラレルワールドの考え方
（同じ出来事でも）





【確率的には同じこと】

まとめ

我が国に降る雨に対して「未曾有の豪雨」、「観測史上初」と聞くと、近年、異常気象が多発しているように感じる

しかし、１００年間の統計を見ると、そのような豪雨も、確率理論（極値統計学；Gunbel分布）が予測するようにしか起きていない
（想定外の出来事も、広い視野に立てば想定内の出来事）

100年に一度、1000年に一度と言った最悪値の推定には、極値統計学（Gumbel分布）が非常に役に立つ

降雨に関して、これまでの100年間が、ほぼ定常確率過程に従っているからといって、今後100年間、それが続くとは限らない（降雨に着目した地球温暖化問題に注視してゆきたい）