利用对数正态分布计算年最大风速

周正强

（上海台风研究所）

提　　要

本文提出用对数正态分布计算年最大风速，这种方法与当前使用的计算方法主要的不同是，它具有较完善的子样检验方法，并且能计算最大风速的容忍上限。

一、引言

在建筑结构设计中，需要按照工程设计要求确定若干年一遇的最大风速，例如50年一遇或100年一遇，甚至更长时间一遇的最大风速，也就是一定概率下的年最大风速。

若把年最大风速视作随机变量ξ，则：

$$R(x)=P(\xi>x)$$

表示风速大于x的概率，上式也就写为：

$$R(x)=1-F(x)$$

其中$F(x)$是最大风速分布。

最大风速的概率计算方法主要有Pearson-Ⅲ型、尺度变换法和Gumbel极值分布[1]—[3]。

二、对数正态分布

50年代，曾应用对数正态分布计算陆地上的年最大风速，但并没有从物理意义给予阐明，同时由于该分布函数的统计特征的表达方式较为复杂，因此该方法在应用上受到限制。

到70年代，许多统计工作者对该分布作了较多的研究工作[4]—[6]，且导出了对数正态分布的机理：当某个随机变量受到许多独立的随机变量"均匀"影响时，该变量是服从对数正态分布的。并简化了该函数的计算方法，建立了一套效率较高的分布拟合方法。

气象上的最大风速，尤其是海上最大风速主要由台风、气旋和冷空气引起，例如东海海域几乎每年受台风影响，最大风速主要由台风引起，台风引起的最大风速可以看作是许多随机因子的"均匀"影响所致，因此，可以用对数正态分布拟合海上年最大风速。

对数正态分布密度函数为：

$$f(y)=\frac{1}{\sqrt{2\pi}\sigma y}\exp[-(\ln y-\mu)^2/2\sigma^2]$$

$$y>0,\quad \sigma>0,\quad -\infty<\mu<+\infty$$

其分布函数为：

$$F(x)=\frac{1}{\sqrt{2\pi}\sigma}\int_0^x\frac{1}{y}\exp[-(\ln y-\mu)^2]/2\sigma^2 dy \qquad (1)$$

可以推得数学期望：

$$E(x)=\exp[(\mu+\sigma^2)/2]$$

方差：

$$D(x)=\exp(2\mu+\sigma^2)\cdot(\exp\sigma^2-1)$$

对函数$F(x)$作更进一步简化：令$x=(\ln y-\mu)/\sigma$，（1）式可写为：

$$F(x)=\frac{1}{2\pi}\int_{-\infty}^{(\ln x-\mu)/\sigma}\exp(-x^2/2)dx \qquad (2)$$

（2）式表明，若随机变量ξ服从对数正态分布，则$\ln\xi\sim N(\mu,\sigma^2)$，于是可记：

$$F(x)=\Phi\left(\frac{(\ln x-\mu)}{\sigma}\right)=P \qquad (3)$$

三、对数正态分布的使用方法

1. 对数正态分布的拟合检验

首先验证年最大风速的子样来自对数正态母体，可以用图估计法或数值法检验。图

估计法比较直观，方法简便，只要把观测序列[1)]一一点绘在一张对数正态概率纸上。若点的分布在纸上呈一直线，说明子样来自对数正态母体，反之则可拒绝。另外在概率纸上可读出参数 μ 和 σ^2 的数值。

数值法检验的方法是对实测数据（最大风速）x 作变换：

$$t_i = \ln x_i \quad (i = 1, 2 \cdots\cdots n) \qquad (4)$$

则对数正态分布的检验问题就转为正态分布的检验。正态分布的检验方法有W（Wilk-Shapiro）检验或D'Agostino检验及偏度检验和峰度检验。它们分别是根据子样分布的特征，检验子样是否来自正态母体，这种检验方法的效率较高，也较常用。具体检验方法可参见《国家标准》（GB）。

2.年最大风速的估算

当子样服从对数正态分布，估算若干年一遇的年最大风速方法有两种。a：在概率纸上直接读数。

b：把实测数据作变换，由（4）式可得到

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}\ln x_i = \frac{1}{n}\sum_{i=1}^{n} t_i = \bar{t};$$
$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(t_i - \bar{t}) \qquad (5)$$

$\hat{\mu}$，$\hat{\sigma}^2$ 分别是函数F(x)中的参数μ、σ^2 的一个估计。

根据（3）式给定P，$P = 1 - R$，($0 < R < 1$)，按正态分布原理，直接在标准正态分布表中查得分位点U_p，

$$U_p = (\ln x - \mu)/\sigma \qquad (6)$$

把（5）式中的 $\hat{\mu}$，$\sqrt{\hat{\sigma}^2}$ 代入（6）式，解出 $x = \hat{x}_p = \exp(U_p \cdot \hat{\sigma} + \mu)$。

$\hat{x}_p$ 是$\frac{1}{R}$年一遇的年最大风速估算值。

3.年最大风速的容忍上限

已知μ、σ，可直接从（3）式计算$\frac{1}{R}$年一遇或 $P = 1 - R$ 概率下的年最大风速x_p。通常μ、σ 是未知的，（5）式中的$\hat{\mu}$、$\hat{\sigma}^2$ 仅由子样估计得到，由此得到的风速也仅是一个点估计，不同的子样可以有不同的点估计$\hat{x}_p$，现设法找一个x_u，让其在给定信度 $1-\alpha$下满足：

$$P_{\mu\sigma}(x_u \geqslant x_p) = 1 - \alpha \qquad (7)$$

x_p 是若干年一遇的风速值，x_u 是 x_p 的上限。换句话说，在给定信度下，保证$\frac{1}{R}$年一遇的年最大风速不超过x_u。这就是容忍上限问题。

对变量作变换，并令 $t = \ln x$，$t \sim N(\mu, \sigma^2)$。并记 $t_u = \ln x_u$，$t_p = \ln x_p$　（8）

把（7）式中的x_u、x_p分别换成t_u、t_p，（7）式可改写为：

$$P_{\mu\sigma}(x_u \geqslant x_p) = P_{\mu\sigma}(t_u \geqslant t_p) = 1 - \alpha \qquad (9)$$

t_u 是 t_p 的上限，由（6）可知，当μ、σ已知时，$U_p = (\ln x_p - \mu)/\sigma = (t_p - \mu)/\sigma$，并有

$$t_p = \mu + \sigma \cdot U_p \qquad (10)$$

若μ，σ未知，则用 $\bar{t} = \frac{1}{n}\sum_{i=1}^{n} t_i$，和 $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(t_i - \bar{t})^2$ 来估计 μ, σ^2，因此，设

$$t_u = \bar{t} + S \cdot K_p \qquad (11)$$

是很自然的，（10）式、（11）式代入（9）式中的第二个等号左端，有：

$P_{\mu,\sigma}(\bar{t} + S \cdot K_p \geqslant t_p) = P_{\mu,\sigma}(\bar{t} + S \cdot K_p \geqslant \mu + \sigma \cdot U_p) = 1 - \alpha$，或有：

$$P_{\mu,\sigma}\left(\frac{\bar{t} - \mu - \sigma \cdot U_p}{S} \geqslant -K_p\right) =$$
$$P_{\mu,\sigma}\left(\frac{\bar{t} - \mu - \sigma \cdot U_p}{S} \cdot \sqrt{n} \geqslant -\sqrt{n}K_p\right)$$
$$= 1 - \alpha \qquad (12)$$

（12）中统计量

1）观测序列指一组逐年最大风速值。

$$\gamma^{\circ}=\frac{(\bar{t}-\mu-\sigma\cdot U_p)}{S}\cdot\sqrt{n}$$

$$=\frac{\frac{(\bar{t}-\mu)}{\sigma}\cdot\sqrt{n}-U_p\cdot\sqrt{n}}{\frac{S}{\sigma}}$$

正是服从中心度$\delta=U_p\cdot\sqrt{n}$，自由度为n－1的非中心t分布，因此，$\bar{t}+s\cdot K_P$是t_P的置信度为1—α的容忍上限。

一旦给定P＝1－R、n和1－α，便可求出$K_p=K_p(n, P, 1-\alpha)$。ISO（国际标准化组织）给出了求K_p的表[7]。把K_p、$\bar{t}$、S代入(11)式，便求出t_u，e^{t_u}便是容忍上限（最大风速），即有(1－α)的把握说$\frac{1}{R}$年一遇的年最大风速不超过$x_u=e^{t_u}m\cdot s^{-1}$。

三、实例计算

计算50年一遇的年最大风速。

已知：海上某地36年的逐年最大风速资料$\{x_i\}$，i＝1，2…36。

计算步骤如下。

1.验证子样是来自对数正态分布母体

用图估计法（图略）或数值检验（方法见《国家标准》）：表明子样来自母体。该地的年最大风速可用对数正态分布计算。

2.计算50年一遇的年最大风速

由（3）式得到：$F(x)=\Phi\left(\frac{\ln x-\mu}{\sigma}\right)=P$，给定$R=\frac{1}{50}=0.02$，$P=1-R=0.98$，查正态分布表得：$U_p=U_{0.98}=2.06$，并解出$U_p=(\ln x-\mu)/\sigma$中的x，式中的μ、σ以$\hat{\mu}=3.394$，$\sqrt{\hat{\sigma}^2}=0.2872$代替，得$x_{0.98}=53.9m\cdot s^{-1}$。即50年一遇的年最大风速为$53.9m\cdot s^{-1}$。

3.计算50年一遇年最大风速的容忍上限

按公式$t_u=\bar{t}+S\cdot K_P$计算（$\bar{t}=\hat{\mu}$，$s=\sqrt{\hat{\sigma}^2}$）n＝36，P＝0.98，1－α＝0.90，查表[8]得$K_p(36, 0.98, 0.90)=2.52$。把K_P、S，$\bar{t}$代入（11）得$t_u=4.11774$，$x_u=\exp t_u=61m\cdot s^{-1}$。

即有90％的把握说50年一遇的年最大风速不会超过$61m\cdot s^{-1}$。

4.根据海上两个石油钻井平台所在位置(29.2°N，125.5°E)、(27.0°N，122.0°E)的风速资料，用该方法作了50年一遇的年最大风速计算，得到有益结果，并把计算的结果提供给有关部门。

还对北京地区20年的风速资料作了30年一遇的年最大风速计算，结果也是满意的(见表1和表2)。表2表明，对于陆地上的风速，对数正态分布拟合理论分布的误差有0.08，与Pearson-Ⅲ相同，但海上风速的拟合较好，误差仅0.03。似乎用对数正态分布拟合海上风速要比陆上的要好些。

表1 计算结果

观测点	参数估计		计算年最大风速（$m\cdot s^{-1}$）				
	$\hat{\mu}$	$\hat{\sigma}^2$	Pearson-Ⅲ	Gumbel	尺度变换	对数正态	对数正态上限
北京地区	2.9362	0.01839	24.0	24.5	25.0	24.2	27.4
2号平台	3.394	0.0825	60.6	57.2	62.5	53.9	61.0
3号平台	3.249	0.0728	49.6	47.6	50.9	45.0	51.0

四、小结

关于年最大风速的统计方法有多种，各种方法计算的结果都是一种较长时间的“预测”，因此难以比较各种方法的优劣。如从计算方法的合理性来看，对数正态分布有以下特点。

1.只要对变量作变换，就可用熟知的正

表 2　理论分布与经验分布差 $|Fn(x) - F(x)|$

方法 δ / 观测点	Pearson-Ⅲ	Gumbel	尺度变换	对数正态
北京地区	0.08	0.048	0.053	0.08
2号平台	0.08	0.05	0.06	0.03
3号平台	0.14	0.11	0.12	0.10

态分布性质加以讨论，并由此给计算过程带来方便。

2. 除图估计法外，尚有效率较高的分布拟合方法——W检验、偏度和峰度检验。从统计意义上说这样的检验方法比较完善。

3. 用统计方法计算最大风速，存在随机误差，希望有一个给定概率下的上限风速，这对工程设计部门是有参考价值，用对数正态分布可以定量给出风速上限。

当然，只有当子样与该分布有较好的拟合，才能体现上述优点，目前仅是一种尝试，有待于更多实例的验证。

参考文献

〔1〕上海数学分会概率论数理统计小组，关于最大风速的数理统计，风压论文选辑，上海科学出版社，1959年。
〔2〕朱瑞兆，风压计算的研究，科学出版社，1979年。
〔3〕朱瑞兆，我国不同的风压计算，气象学报，1984年第2期。
〔4〕林少宫，概率论与数理统计，人民出版社，1980年。
〔5〕茆诗松、王玲玲，可靠性统计，华东师范大学出版社，1984年。
〔6〕迈耶（美），科技工作中的数据分析（中译本），原子能出版社，1984年。
〔7〕四川环保科研监测所（译），ISO数理统计方法标准译文集，四川科学出版社，1984年。
〔8〕戴树理等，可靠性试验及其统计分析，国防工业出版社，1983年。

Calculation of annual maximum wind speed using logarithmic normal distribution

Zhou Zhengqiang

(Shanghai typhoon institute)

In this paper, logarithmic normal distribution is used to calculate annual maximum wind speed. Main differences of this method from others are that there is a more perfect subsample inspection and the upper limit of tolerance can be calculated.