

我国地面气温极值和地面风速极值的渐近分布*

曲延禄 阎书源 张程道

(总参气象研究所)

提 要

利用 1951—1982 年期间我国 173 个测站的地面最高气温年极值记录, 158 个测站的地面最低气温年极值记录和 83 个测站的地面最大风速年极值记录, 通过统计推断, 找出了我国地面最高、最低气温年极值和地面最大风速年极值遵循的渐近分布——Weibull 分布和 Gumbel 分布, 并简要讨论了它们的形状。

一、引 言

气候要素极值分布在气候统计学、经济建设和国防建设中有重要应用。气候要素极值作为气候随机变量在数学意义上是不稳定的, 但它们随时间的变化(气候随机变化)过程在概率意义上却可以是稳定的^[1]。因此, 气候要素极值的分布有可能用概率模式(分布函数)去模拟。Gumbel 分布就是这样一种被用来拟合气候极值记录的 概率模式^[2]。但是, 它不能对任何气候要素极值、在任何地区都给出满意的拟合结果。本文利用 1951—1982 年我国部分测站的地面最高、最低气温和地面最大风速极值记录, 分别用 Gumbel 分布函数和 Weibull 分布函数进行了拟合试验, 证明: 对于给定的气候要素极值, 不同的测站并非都遵循同一分布。在此基础上, 找出了我国地面最高、最低气温极值和最大风速极值服从的渐近分布函数。

二、分布函数模式和参数估计方法

本文用于拟合地面最高、最低气温极值和最大风速极值记录的分布函数采用 Gumbel 分布和 Weibull 分布。Gumbel 分布是极值渐近分布的一种理论模式, 其分布函数和密度函数分别为

$$F(x) = P(\xi < x) = e^{-e^{-A(x-B)}} \quad (1)$$

$$f(x) = Ae^{-A(x-B)-e^{-A(x-B)}} \quad (-\infty < x < \infty) \quad (2)$$

对于极小值, Gumbel 分布函数和密度函数具有如下形式

* 本文于 1985 年 10 月 5 日收到, 1987 年 6 月 8 日收到最后修改稿。

$$F_1(x) = P(\xi < x) = 1 - e^{-A_1(x-B_1)} \quad (3)$$

和

$$f_1(x) = \frac{dF_1(x)}{dx} = A_1 e^{A_1(x-B_1)} - e^{A_1(x-B_1)} \quad (4)$$

式中的参数A和B(或 A_1 和 B_1)容易由观测记录求出。

Weibull分布的分布函数和密度函数分别为

$$F(x) = P(\xi < x) = 1 - e^{-\frac{(x-\lambda_3)^{\lambda_2}}{\lambda_1}} \quad (5)$$

和

$$f(x) \begin{cases} = \frac{\lambda_2}{\lambda_1} (x-\lambda_3)^{\lambda_2-1} e^{-\frac{(x-\lambda_3)^{\lambda_2}}{\lambda_1}}, & x > \lambda_3 \\ = 0, & x \leq \lambda_3 \end{cases} \quad (6)$$

式中,

$$F(x) = 1 - \frac{1}{T(x)}, \quad T(x_i) = \frac{N+1}{j} \quad (7)$$

$\lambda_1 (> 0)$ 为尺度参数, $\lambda_2 (> 0)$ 为形状参数, λ_3 为位置参数。当计算极小值的分布时, (5) 和 (6) 式应改为

$$F_1(x) = P(\xi < x) = e^{-\frac{(-x+\lambda_3^*)^{\lambda_2}}{\lambda_1}} \quad (8)$$

和

$$f_1(x) \begin{cases} = \frac{\lambda_2}{\lambda_1} (-x+\lambda_3^*)^{\lambda_2-1} e^{-\frac{(-x+\lambda_3^*)^{\lambda_2}}{\lambda_1}}, & x < \lambda_3^* \\ = 0, & x \geq \lambda_3^* \end{cases} \quad (9)$$

式中, $\lambda_3^* = -\lambda_3$, $F_1(x)$ 和 $T(x_i)$ 的表达式仍为(7)式。三参数 λ_1, λ_2 和 λ_3 (或 λ_3^*) 用最大似然法估计(详见文献[3])。

三、我国地面气温极值和地面风速极值遵循的渐近分布

我们采用统计推断方法来确定气候要素极值遵循的渐近分布函数。由于总体分布函数的形式是未知的, 我们直接用某些已知的分布函数拟合气候要素极值记录, 通过比较、评定拟合标准差, 以及拟合(理论分布)分位数与观测(频数分布)分位数间的偏差, 从中选择出拟合标准差小, 而且拟合分位数与观测分位数偏差也小的分布模式作为假设分布, 然后通过对分布进行适度测验, 最后确定出气候要素极值应遵循的渐近分布函数。

1. 地面气温年极值的渐近分布

我们选定 Gumbel 分布和 Weibull 分布作为已知分布形式, 分别对 1951—1982 年我国 173 个测站的地面最高气温和 158 个测站的地面最低气温年极值记录进行了拟合。结果表明, 对于最高气温年极值, Weibull 分布的拟合标准差 σ_w 小于 Gumbel 分布的拟合标准差 σ_g 的测站数为 102 个, $\sigma_g < \sigma_w$ 的测站数为 71 个, 二者之比为 $\frac{102}{71} \approx \frac{1.44}{1}$; 对于最低气温年极值, $\sigma_w < \sigma_g$ 的测站数为 105 个, $\sigma_g < \sigma_w$ 的测站数为 53 个, 二者之

比为 $\frac{105}{53} = \frac{1.98}{1}$ 。我们还计算了各测站的 50%, 75%, 90%, 95% 和 97% 分位数的拟合误差(拟合值与由频数分布求得的对值之差,下同)。表 1 和表 2 列出的是最高气温年极值 $\sigma_w < \sigma_c$ 和 $\sigma_c < \sigma_w$ 的各 30 个测站的两种分布的拟合标准差和分位数的拟合误差。可以看出,对于最高气温年极值, $\sigma_w < \sigma_c$ 的各测站 Weibull 分布的分位数拟合误差的绝对值平均小于 Gumbel 分布的误差; $\sigma_c < \sigma_w$ 的各测站的 Gumbel 分布的分位数拟合误差的绝对值平均小于 Weibull 分布的误差。从统计误差的符号可以看出,大多数的误差为负值。这表明无论是用 Weibull 分布,还是用 Gumbel 分布拟合最高气温年极值,拟合值大多稍偏低于实际出现值。

用 Колмогоров 适度测验进行检验,频数分布与 Weibull 分布或者与 Gumbel 分布的差异均是不显著的(取 $Q(\lambda) = \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2 \lambda^2} = 0.95$)。而且一般地说,拟合标准差越小的分布,这种差异越不显著。

综上,我们有理由对指定的测站在 Weibull 和 Gumbel 两种分布中选择拟合标准差小的分布,作为该站最高气温年极值的渐近分布。根据 173 个测站的统计结果,我国最高气温年极值的渐近分布,约有 59% 的测站属于 Weibull 分布,约有 41% 的测站为 Gumbel 分布。

对于最低气温年极值,拟合标准差小的分布,其分位数(特别是 95% 以上的高分位数)的拟合误差平均地说也小些(计算结果从略)。与最高气温年极值不同的是,最低气温年极值的分位数拟合误差大多为正值。这意味着,由选用的两种分布拟合出的最低气温年极值大多偏高于实际出现值。通过比较还可看出,最低气温年极值各分位数的拟合误差平均大于最高气温年极值对应分位数的拟合误差。这反映出最低气温年极值的随机变化过程的概率稳定性不如最高气温年极值的强。

用 Колмогоров 适度测验对最低气温年极值进行检验的结果同于最高气温年极值。因此,对于最低气温年极值,同样有理由在 Weibull 和 Gumbel 两种分布中,选择拟合标准差小的分布作为渐近分布。根据对 158 个测站的统计,我国最低气温年极值的渐近分布,约有 66% 的测站属于 Weibull 分布,约有 34% 的测站属于 Gumbel 分布。

2. 地面最大风速年极值的渐近分布

地面最大风速系指按 10min 间隔平均求得的最大风速。我们同样以 Gumbel 和 Weibull 分布作为已知分布形式,分别对 1951—1982 年我国 83 个测站的地面最大风速年极值记录进行了拟合(计算结果从略)。在这些测站中, $\sigma_w < \sigma_c$ 的测站数与 $\sigma_c < \sigma_w$ 的测站数之比为 $\frac{43}{40} \approx \frac{1.1}{1}$ 。地面最大风速年极值的拟合标准差和分位数的拟合误差均比气温年极值的来得大。其原因是目前业务用的风速测定仪器的误差大于温度测定仪器的误差;其次,在采样期间,我国气象台站的风速测定仪器曾三次换装^[4],使得风速测定误差变得更加复杂。但无论是 $\sigma_w < \sigma_c$ 的测站,还是 $\sigma_c < \sigma_w$ 的测站,两种分布的分位数平均拟合误差的差值均不超过 0.6m/s,这种差异应当看成是不明显的。Колмогоров 适度测验的结果表明,频数

表 1 年最高气温的拟合标准差 $\sigma_w < \sigma_g$ 的 30 个测站的两种分布的拟合标准差和分位数的拟合误差 ($^{\circ}\text{C}$)

站 名	Weibull 分 布						Gumbel 分 布					
	拟合标准差 σ_w	50%	75%	90%	95%	97%	拟合标准差 σ_g	50%	75%	90%	95%	97%
嫩 江	0.225	-0.2	-0.1	0.0	-0.6	-0.6	0.406	-0.6	-0.3	0.3	0.4	0.7
哈 尔 滨	0.204	-0.1	-0.2	0.0	-0.3	0.5	0.466	-0.4	-0.3	0.4	1.1	1.7
丹 东	0.164	0.1	-0.2	-0.3	-0.1	0.0	0.282	-0.2	-0.3	0.0	0.6	1.0
大 连	0.263	-0.1	-0.4	-0.1	0.3	-0.1	0.265	-0.1	-0.4	1.1	0.8	0.5
乌 鲁 木 齐	0.351	-0.2	-0.4	-0.7	-0.5	-0.2	0.543	-0.6	-0.6	-0.2	0.6	1.5
吐 鲁 番	0.289	0.1	0.0	-0.2	0.2	0.4	0.357	0.0	0.0	0.1	0.8	1.3
喀 什	0.370	0.2	-0.4	-0.3	0.0	0.2	0.473	0.0	-0.5	0.1	0.9	1.6
延 安	0.357	-0.1	0.0	0.2	-0.9	-0.7	0.371	-0.3	-0.2	0.4	-0.3	-0.4
西 安	0.123	0.0	-0.1	-0.4	-0.5	-0.3	0.198	-0.2	-0.2	-0.2	0.1	0.5
锡林浩特	0.178	-0.2	-0.3	0.0	-0.3	-0.2	0.497	-0.5	-0.5	0.4	0.7	1.3
呼 和 浩 特	0.229	0.0	-0.4	-0.5	-0.3	0.0	0.318	-0.2	-0.5	-0.1	0.5	1.2
北 京	0.237	-0.3	-0.4	-0.7	-0.4	-0.3	0.425	-0.6	-0.5	-0.2	0.7	1.3
天 津	0.221	-0.1	-0.5	-0.4	-0.2	0.0	0.438	-0.5	-0.6	-0.1	0.7	0.4
青 岛	0.341	-0.1	-0.5	-0.5	-0.1	0.2	0.382	-0.1	-0.5	-0.2	0.4	1.0
徐 州	0.128	-0.1	-0.5	-0.4	-0.3	-0.1	0.458	-0.5	-0.6	0.1	0.7	1.4
蚌 埠	0.264	-0.1	-0.4	-0.4	-0.3	0.0	0.318	-0.2	-0.4	-0.2	0.3	1.0
杭 州	0.164	0.0	-0.2	-0.2	-0.2	-0.1	0.231	-0.2	-0.2	0.0	0.4	0.8
厦 门	0.136	-0.1	-0.2	-0.3	-0.3	-0.1	0.261	-0.3	-0.3	0.0	0.4	0.7
安 阳	0.217	0.0	0.0	-0.1	-0.2	-0.2	0.400	-0.3	-0.2	0.2	0.4	0.7
郑 州	0.241	-0.2	-0.6	-0.8	-0.6	-0.3	0.284	-0.3	-0.7	-0.5	0.1	0.6
武 汉	0.162	-0.2	-0.2	-0.3	-0.3	0.0	0.180	-0.2	-0.1	-0.2	0.0	0.5
广 州	0.166	0.0	-0.2	-0.6	-0.7	-0.5	0.187	-0.1	-0.3	-0.4	-0.2	0.1
海 口	0.128	-0.1	-0.3	-0.2	-0.1	0.0	0.277	-0.3	-0.4	0.0	0.4	-0.2
梧 州	0.138	-0.2	-0.2	-0.3	-0.2	0.0	0.180	-0.3	-0.2	-0.1	0.2	0.6
南 充	0.185	-0.2	-0.3	-0.5	-0.6	-0.7	0.257	-0.5	-0.4	0.1	0.1	0.3
遵 义	0.217	-0.2	-0.4	-0.2	-0.3	-0.2	0.230	-0.2	-0.4	-0.1	0.1	0.4
贵 阳	0.201	-0.4	-0.3	-0.3	-0.5	-0.5	0.241	-0.2	-0.2	-0.2	-0.5	-0.5
兴 仁	0.184	-0.2	-0.4	-0.7	-0.6	-0.4	0.245	-0.4	-0.5	-0.4	0.1	0.5
腾 冲	0.117	-0.1	-0.3	0.0	0.1	0.1	0.235	-0.3	-0.4	0.1	0.6	0.8
昆 明	0.159	-0.1	-0.3	-0.2	0.1	0.2	0.339	-0.4	-0.5	0.1	0.7	1.1
误差绝对值的平均	0.212	0.13	0.29	0.33	0.34	0.24	0.325	0.30	0.37	0.22	0.46	0.82
负误差个数 正误差个数	—	22/8	27/3	25/5	25/5	17/13	—	28/2	29/1	14/16	3/27	3/27

表 2 年最高气温的拟合标准差 $\sigma_a < \sigma_w$ 的 30 个测站的两种分布的拟合标准差和分位数的拟合误差 ($^{\circ}\text{C}$)

站 名	Gumbel 分 布						Weibull 分 布					
	拟合标准差 σ_a	50%	75%	90%	95%	97%	拟合标准差 σ_w	50%	75%	90%	95%	97%
长 春	0.251	-0.1	0.0	-0.2	-0.1	0.1	0.272	-0.2	-0.1	-0.4	-0.5	-0.4
沈 阳	0.289	-0.4	0.0	-0.3	-0.7	-0.4	0.333	-0.4	-0.1	-0.5	-1.1	-1.0
库 车	0.144	-0.1	-0.1	-0.5	-0.2	0.3	0.188	-0.1	-0.2	-0.7	-0.4	-0.1
西 宁	0.194	-0.2	0.0	-0.3	-0.4	-0.3	0.234	-0.2	0.0	-0.4	-0.7	-0.6
敦 煌	0.265	-0.4	-0.5	-0.5	-0.7	-0.3	0.299	-0.3	-0.5	-0.8	-1.2	-1.1
酒 泉	0.250	-0.2	0.0	0.2	-0.3	-0.2	0.287	-0.3	0.0	0.1	-0.5	-0.5
兰 州	0.284	-0.3	-0.3	-0.8	-0.8	-0.5	0.296	-0.2	-0.3	-1.1	-1.4	-1.3
银 川	0.235	-0.3	0.1	0.3	-0.5	-0.7	0.282	-0.3	0.1	0.2	-0.8	-1.2
太 原	0.270	-0.1	0.0	-0.6	-0.8	-0.5	0.328	-0.2	0.0	-0.7	-1.1	-0.9
承 德	0.474	-0.3	0.0	-0.3	-0.5	-0.1	0.556	-0.3	0.0	-0.6	-1.1	-0.9
石 家 庄	0.168	-0.1	-0.2	-0.3	0.0	0.3	0.211	-0.2	-0.2	-0.5	-0.4	-0.2
济 南	0.184	-0.2	-0.2	-0.1	-0.2	-0.2	0.208	-0.3	-0.3	-0.2	-0.5	-0.7
南 京	0.216	-0.1	-0.1	-0.3	0.1	0.5	0.241	-0.1	-0.1	-0.4	-0.3	0.0
南 通	0.079	-0.2	-0.2	-0.4	-0.4	-0.1	0.120	-0.3	-0.3	-0.5	-0.5	-0.3
合 肥	0.145	-0.2	-0.3	-0.5	-0.3	0.1	0.202	-0.2	-0.3	-0.6	-0.6	0.5
上 海	0.096	-0.2	-0.4	-0.6	-0.3	0.0	0.122	-0.3	-0.5	-0.7	-0.5	-0.2
温 州	0.145	-0.3	-0.2	-0.6	-0.3	0.1	0.164	-0.3	-0.2	-0.7	-0.7	-0.4
福 州	0.152	-0.1	0.0	-0.1	0.1	0.5	0.157	-0.2	0.0	0.3	-0.1	-0.2
南 昌	0.133	0.0	-0.1	0.1	-0.1	0.0	0.140	0.0	-0.1	0.0	-0.5	-0.4
信 阳	0.228	-0.0	-0.5	-0.5	0.0	0.6	0.234	0.0	-0.5	-0.7	-0.4	0.2
长 沙	0.109	-0.3	-0.4	-0.3	-0.1	0.3	0.116	-0.3	-0.4	-0.5	-0.3	-0.1
芷 江	0.147	-0.3	-0.8	-0.7	-0.2	-0.2	0.160	-0.4	-0.9	-0.8	-0.4	0.0
韶 关	0.244	-0.2	-0.1	-0.2	-0.7	-0.7	0.273	-0.2	-0.2	-0.4	-1.0	-1.1
汕 头	0.145	-0.1	-0.2	-0.6	-0.4	-0.1	0.174	-0.3	-0.2	-0.6	-0.3	0.0
桂 林	0.128	0.0	-0.3	-0.3	-0.2	0.0	0.133	0.0	-0.4	-0.4	-0.3	-0.2
南 宁	0.148	-0.2	-0.2	-0.2	-0.2	-0.3	0.183	-0.3	-0.3	-0.3	-0.5	-0.5
松 潘	0.146	-0.0	-0.5	-0.3	-0.1	0.2	0.175	-0.1	-0.6	-0.5	-0.4	-0.2
成 都	0.128	0.1	0.1	-0.5	-0.4	-0.1	0.167	0.1	0.0	-0.6	-0.6	-0.4
重 庆	0.158	-0.1	-0.4	0.2	-0.1	0.0	0.180	-0.1	-0.4	0.0	-0.5	-0.6
拉 萨	0.223	-0.2	-0.3	-0.5	-0.5	-0.3	0.245	-0.2	-0.3	-0.7	-0.9	-0.9
误差绝对值的平均	0.193	0.18	0.22	0.38	0.32	0.27	0.223	0.21	0.25	0.50	0.62	0.50
负误差个数 正误差个数	—	27/3	21/9	26/4	26/4	16/14	—	26/4	23/7	25/5	30/0	26/4

分布与Weibull分布或与Gumbel分布的差异也都是不显著的。因此,可取这两种分布的任何一种作为地面最大风速年极值的渐近分布。但是,考虑到97%高分位数的拟合误差,Weibull分布要比Gumbel分布稍小些,为了减少采用几种分布模式带来的麻烦,不妨只取Weibull分布作为我国地面最大风速年极值的渐近分布函数。用Weibull分布拟合的分位数的负误差的次数占优势,因此用这种分布拟合出的风速值大多偏小些。

四、渐近分布函数的形状

渐近分布函数的形状是由分布的参数决定的。作为例子,在图1和图2中分别绘出了以Weibull分布和Gumbel分布为渐近分布函数的天津和南京两站的最高气温年极值的拟合曲线及与之对应的密度函数曲线。

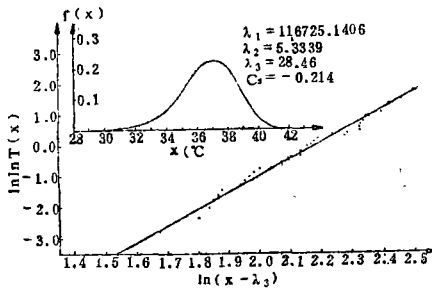


图1 天津最高气温年极值的Weibull分布的拟合曲线和密度函数曲线 (图中的圆点为观测值,资料年数为32a)

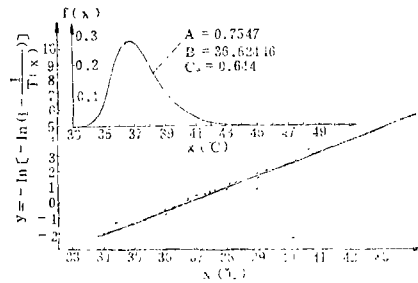


图2 南京最高气温年极值的Gumbel分布的拟合曲线和密度函数曲线 (图例说明同图1)

由(6)或(9)式可见,Weibull分布的密度曲线的形状是受 λ_1 , λ_2 和 λ_3 (或 λ_3^*)三个参数控制的。 λ_3 是密度曲线的左端点坐标(见图1),其数值与随机变量的最小值(对于极大值的渐近分布)有关; λ_3^* 是密度曲线的右端点坐标,其数值与随机变量的最大值(对于极小值的渐近分布)有关, λ_1 控制着密度曲线在横轴方向上伸展的宽度。 λ_2 决定了曲线的形状,是三参数中最重要的一个参数。不同测站间的 λ_2 值的变化比较平稳,而 λ_1 值的变化却是十分剧烈的。这正是显示了 λ_2 参数在决定分布形状上的重要作用。有关Gumbel分布的形状在许多文献(例如[5])中已有讨论,此处不再赘述。

对偏倚系数 C_s 的计算结果表明,气温年极值渐近分布的偏倚系数(绝对值)平均小于风速年极值的偏倚系数(绝对值)。这意味着风速年极值渐近分布较之气温年极值分布对正态分布有着更大程度的偏离。在气温极值的两种渐近分布中,Gumbel分布的偏倚系数(绝对值)平均大于Weibull分布的偏倚系数(绝对值)(频数分布的偏倚系数也是如此)。这种分布形状上的差异,恐怕就是不同测站的气温年极值遵循不同渐近分布的原因。

五、结 论

1. 不同的气候要素极值可以遵循不同的渐近分布,而且不同测站的同一气候要素极

值也可以有不同的渐近分布函数。

2. 我国不同测站的地面气温年极值的渐近分布分别为Weibull分布或Gumbel分布。其中,最高气温年极值分布属于前者的约占 59%,属于后者的约占 41%;最低气温年极值分布属于前者的约占66%,属于后者的约占34%。对于指定的测站,可通过比较上述两种分布的拟合标准差来判断气温年极值所属的分布类型。地面最大风速年极值遵循Weibull分布。

3. 地面最大风速年极值渐近分布的偏倚系数(绝对值)平均大于气温年极值渐近分布的偏倚系数(绝对值)。气温年极值的Gumbel分布的偏倚系数(绝对值)平均大于Weibull分布的偏倚系数(绝对值)。

参 考 文 献

- [1] 么枕生, 气候统计学的研究展望, 气象科技, 1984, 6.
- [2] Sissenwine, N., et al. Synopsis of background material for MIL-STD-210B, Climatic extremes for military equipment, AFCL, L. G. Hanscom Field, Massachusetts, AD-780 508, 1974.
- [3] 曲延禄、张程道、阎书源, 三参数 Weibull 分布的参数估计, 气象学报, 45, 3, 374—378, 1987.
- [4] 中央气象局编定, 地面气象观测规范, 178, 气象出版社, 1979.
- [5] 屠其璞等, 气象应用概率统计学, 210—216, 气象出版社, 1984.

ON ASYMPTOTIC DISTRIBUTIONS FOR EXTREMES OF SURFACE TEMPERATURE AND SURFACE WIND OVER CHINA

Qu Yanlu Yan Shuyuan Zhang Chengdao

(*Meteorological Research Institute, Headquarters of the General Staff*)

Abstract

According to statistic infer method, the asymptotic distributions, Weibull distribution or Gumbel distribution, for yearly extremes of surface temperature and surface wind in China were discovered. In this study we used the data of 173 stations for yearly maximum surface temperature and 158 stations for yearly minimum surface temperature and 83 stations for yearly maximum surface wind during the periods from 1951 to 1982. Finally, in this paper, the characteristics of the asymptotic distributions were discussed briefly.