

瑞安市暴雨强度概率分布公式参数估计研究

顾骏强 陈海燕 徐集云

(浙江省气候中心, 杭州 310017)

提 要

该文根据瑞安市 1965~1997 年降水自记记录, 通过适当的概率模式对暴雨强度抽样频数分布进行拟合, 并由此编制反映暴雨强度、历时、重现期三者关系的 $i-t-T$ 表. 利用最优化法进行非线性超定方程组的求解, 从而实现对暴雨强度公式中参数的估计.

关键词: 概率模式 参数估计 暴雨强度公式

引 言

根据《室外排水设计规范(GBJ14-87)》^[1]规定, 在进行城市排水管网设计时, 雨水管网和合流制管网的设计排水量, 均应通过本地的城市暴雨强度公式计算, 排水工程的可靠性与采用的暴雨强度公式有直接的关系, 暴雨强度公式是城市雨水排水设计流量的根本依据.

瑞安市位于浙江省东南沿海地区, 地处亚热带季风气候区, 暴雨是该市主要气候灾害之一. 据统计, 瑞安市平均每年暴雨日数 5.7 d. 近年来, 随着沿海经济的腾飞, 瑞安市市政建设迅猛发展, 科学地规划城市发展, 合理地设计市政工程, 是建设部门面临的实际问题. 为了配合市政建设的需要, 使市政工程建立在科学的基础上, 本文对瑞安市暴雨强度概率分布及其强度公式进行了研究, 为该市建立科学的城市雨、污分流的管渠系统提供可靠的依据.

1 瑞安市暴雨抽样资料的采样

瑞安市气象站成立于 1960 年, 位于瑞安市城镇范围之内, 配有自记雨量计, 常年进行雨量观测, 迄今共积累了 38 年的自记雨量记录. 本文根据《室外排水设计规范》^[1]和《给水排水设计手册》^[2]中有关编制暴雨强度公式规定, 从瑞安市气象局 1965~1997 年逐日降水自记曲线上, 按 5、10、15、20、30、45、60、90、120 min 9 个降水历时挑选并摘录历年最大的若干组雨量数据.

在资料采集过程中, 为了不遗漏历史资料中一些大的降雨强度样本, 同时也为了规范资料采集工作, 规定每年各历时挑选最大的 8 组雨量数据. 本文从瑞安市历史降水自记曲线上共摘录 264 组数据, 通过计算机信息化处理, 将选出的资料按从大到小顺序排列(各历时分别进行), 并取前 132 组数据(年数的 4 倍)构成暴雨抽样基本数据.

* 本文得到浙江省重大科研项目(991103364)资助.
1999-04-26 收到, 1999-06-29 收到修改稿

2 瑞安市暴雨强度抽样概率分布模式

在气象要素概率分布的应用中,通常根据不同要素特点采用不同的分布模式进行概率曲线的拟合.本文运用耿贝尔(Gumbel)分布、指数分布、韦伯(Weibull)分布、皮尔逊III型分布等概率模式对瑞安市暴雨强度抽样的分布进行讨论,并根据这些分布模式对暴雨强度资料的拟合情况,通过对比分析确定用于描述瑞安市暴雨强度抽样分布的概率模式.

2.1 概率分布模式^[3]

(1) 耿贝尔(Gumbel)分布

耿贝尔分布的概率密度函数和分布函数分别为:

$$f(x) = ae^{-y}e^{-e^{-y}} \quad (1)$$

$$F(x) = e^{-e^{-y}} \quad (2)$$

式中 y 为抽样 $x_i (i=1, 2, \dots, N)$ 的函数, N 为抽样数. 耿贝尔推得: $y = a(x_i - b)$, 其中 $a > 0$, 为尺度参数, b 是分布密度的众数. 通过积分, 可以求得 y 的数学期望值和均方差分别为:

$$m_y = 0.577216 \quad (3)$$

$$\sigma_y = \pi / \sqrt{6} \quad (4)$$

因此, 由 $y = a(x_i - b)$, 可以得到:

$$a = \frac{\pi}{6} \sigma_x \quad (5)$$

$$b = m_x - \frac{\sigma_x}{\sigma_y} m_y \quad (6)$$

这里 m_x 和 σ_x 分别为 x_i 的数学期望和均方差.

分布函数参数确定以后, 通过下式求对应于重现期 T 的分位数:

$$X_T = b - \frac{1}{a} \ln(-\ln(1 - \frac{1}{T})) \quad (7)$$

(2) 指数分布

指数分布的概率密度函数和分布函数分别为:

$$f(x) = ae^{-a(x-r)} \quad (8)$$

$$F(x) = \int_r^x f(x) dx = 1 - e^{-a(x-r)} \quad (9)$$

式中 $x_i (i=1, 2, \dots, N)$ 为抽样资料, N 为抽样数. a 表示分布曲线离散程度的参数, r 为分布曲线的下限. 故对应于重现期 T 的分位数:

$$X_T = r - \frac{1}{a} \ln T \quad (10)$$

令 $1/a = \alpha$, 利用抽样的经验概率分布和最小二乘法, 可以计算出系数 α 、 r 的估计, 并通过式(10)计算各重现期的分位数.

(3) 皮尔逊(Pearson)III型分布

皮尔逊Ⅲ型分布密度函数和分布函数为:

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} (x - a)^{\alpha-1} e^{-\beta(x-a)} \quad (11)$$

$$F(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_a^x (x - a)^{\alpha-1} e^{-\beta(x-a)} dx \quad (12)$$

式中抽样变量 $x \geq 0$, 参数 $\alpha, \beta \geq 0$ 分别表示概率曲线的形状和尺度参数, a 为概率曲线起点与序列零点的距离. 用皮尔逊Ⅲ型分布拟合暴雨强度抽样资料, 首先要对参数 α, β, a 作估计. 根据矩法得到皮尔逊Ⅲ型的数学期望、方差和偏差系数分别为: $\bar{x} = a + \frac{\alpha}{\beta}$, $\sigma^2 = \frac{\alpha}{\beta^2}$,

$C_s = 2/\alpha$. 由此得到:

$$\begin{cases} \alpha = \frac{4}{C_s^2} \\ \beta = \frac{2}{C_s \sigma} \\ a = \bar{x} - \frac{2\sigma}{C_s} \end{cases} \quad (13)$$

由于高阶矩误差较大, 因此, 矩法对皮尔逊Ⅲ型参数的估计被认为误差偏大. 为了提高估计精度, 可用权函数对参数进行修正, 即:

$$\begin{cases} \hat{C}_s = \frac{4\sigma E}{G} \\ \alpha = \frac{4}{\hat{C}_s^2} \\ \beta = \frac{2}{\hat{C}_s \hat{\sigma}} \\ a = \bar{x} - \frac{2\hat{\sigma}}{\hat{C}_s} \end{cases} \quad (14)$$

其中 \bar{x} 和 $\hat{\sigma}$ 为样本的均值和均方差, 而 $E = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x}) \psi(x_i)$, $G = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \psi(x_i)$, $\psi(x) = \frac{1}{\sigma} e^{-\frac{0.5(x-\bar{x})^2}{\sigma^2}}$. $\psi(x)$ 是正态概率分布密度函数. 参数估计加权后, 减少了由于抽样频数分布尾部取值较少而造成的误差, 提高了计算精度, 这就是所谓的“权函数法”.

皮尔逊Ⅲ型概率分布模式参数确定后, 则可根据达到某一精度要求的拟合曲线进行各概率或重现期对应的分位数的估算. 可以证明皮尔逊Ⅲ型分布曲线某一概率或重现期对应的分位数只依赖于参数 α 或偏度系数 C_s . 因此, 给定一个参数 α 或 C_s 值后, 就可通过曲线离均系数表, 查算出概率 P 或重现期 T 对应的分位数. 若对 C_s 作适当调整, 使拟合误差达到最小, 这便是所谓的“适线法”.

(4) 韦伯(Weibull)分布

三参数韦伯分布的概率密度函数和分布函数为:

$$f(x) = \frac{c}{b} \left(\frac{x-a}{b} \right)^{c-1} e^{-\left(\frac{x-a}{b}\right)^c} \quad x \geq a \quad (15)$$

$$F(x) = 1 - e^{-\left(\frac{x-a}{b}\right)^c} \quad (16)$$

式中 $b, c > 0$ 及 a 分别为尺度参数、形状参数和位置参数。当 $a = 0$ 时, 即为二参数 Weibull 分布。当 $a = 0, c = 1$ 时, 即为指数分布。因此, 指数分布可看作韦伯分布的一个特例。这里选用矩法对参数 a, b, c 作估计。

假设 \bar{x}, σ, C_s 分别为抽样的平均值、均方差和偏度系数。利用前三阶矩, 通过积分可以求得:

$$\begin{cases} \bar{x} = b\Gamma\left(1 + \frac{1}{c}\right) + a \\ \sigma^2 = b^2\left[\Gamma\left(1 + \frac{2}{c}\right) - \Gamma^2\left(1 + \frac{1}{c}\right)\right] \\ C_s = \frac{\Gamma\left(1 + \frac{3}{c}\right) - 3\Gamma\left(1 + \frac{2}{c}\right)\Gamma\left(1 + \frac{1}{c}\right) + 2\Gamma^3\left(1 + \frac{1}{c}\right)}{\left[\Gamma\left(1 + \frac{2}{c}\right) - \Gamma^2\left(1 + \frac{1}{c}\right)\right]^{3/2}} \end{cases} \quad (17)$$

令 $d = \Gamma\left(1 + \frac{2}{c}\right) - \Gamma^2\left(1 + \frac{1}{c}\right)$, $w = \Gamma\left(1 + \frac{1}{c}\right)$ 。则 $b = \sigma / \sqrt{d}$, $a = \bar{x} - bw$ 。利用这些关系式可解出参数 c, b, a 估计值。由于韦伯分布函数 $F(x)$ 与最小值重现期 $T(x)$ 的关系为:

$$T(x) = \frac{1}{1 - F(x)} = e^{\left(\frac{x-a}{b}\right)^c} \quad (18)$$

故重现期 T 所对应的分位数为:

$$X_T = a + b(\ln T)^{1/c} \quad (19)$$

2.2 暴雨强度抽样概率分布

暴雨强度抽样分布曲线的拟合是构造暴雨强度公式的一个重要环节, 过去对暴雨强度抽样的拟合常局限于指数分布和皮尔逊 III 型分布。这里利用前述讨论的各概率分布模式及有关模式参数估计方法对瑞安市暴雨抽样进行拟合, 并通过拟合结果的对比分析来评判各种分布模式的适用性, 最终确定暴雨强度公式编制采用的概率模式。对于拟合效果的评判采用均方根误差、相对均方根误差和绝对值误差等最优化准则, 即:

$$\begin{cases} \sigma = \left(\sum_{i=1}^N (x_i - x_{li})^2 / N \right)^{1/2} \\ f = \sigma / \bar{X} \\ g = \frac{1}{N} \sum_{i=1}^N |X_i - X_{li}| \end{cases} \quad (20)$$

式中 σ, f, g 分别表示均方根误差、相对均方根误差和绝对值误差; X_{li} 为理论概率模式对应于某一观测经验概率 $i/(1+N)$ 的估计值。

表 1 为瑞安市暴雨强度抽样各种概率分布模式时的参数估计。表 2 为各概率分布模式对暴雨强度抽样拟合指标。从各模式的计算结果来看, 耿贝尔分布对各历时拟合效果不佳, 误差偏大, 该分布模式未被用于暴雨强度公式的构造。指数分布、皮尔逊 III 型分布和韦伯分布的拟合情况较好, 均方根误差和相对均方根误差均小于 0.05 mm/min 和 5% 。故

认为指数分布、皮尔逊Ⅲ型分布和韦伯分布合适抽样的描述。其中指数分布和皮尔逊Ⅲ型分布对历时抽样的拟合均方根误差, 平均值小于 0.03 mm/min, 可见这两种分布模式对暴雨强度抽样的拟合达到了很高的精度。

表 1 各种概率分布模式的参数估计

模 式	参 数	历 时 (m in)								
		5	10	15	20	30	45	60	90	120
耿贝尔分布	a	0.5365	0.4380	0.3961	0.3629	0.2980	0.2468	0.2282	0.1939	0.1802
	b	1.7607	1.3550	1.1415	0.9876	0.7839	0.6028	0.4887	0.3610	0.2923
指数分布	γ	3.5528	2.8268	2.4714	2.2049	1.7817	1.4262	1.2521	1.0104	0.8932
	α	0.4427	0.3638	0.3287	0.3009	0.2465	0.2033	0.1886	0.1605	0.1484
韦伯分布	a	1.4697	1.1093	0.9502	0.8005	0.5772	0.4127	0.3303	0.2516	0.2220
	b	0.5037	0.4224	0.3394	0.3274	0.3395	0.3054	0.2601	0.1877	0.1311
	c	1.1771	1.1977	1.0981	1.1444	1.3760	1.4714	1.3763	1.2008	0.9769
皮尔逊Ⅲ分布	α	1.5379	1.5822	1.1764	1.3607	2.4605	3.1442	2.4531	1.5727	0.8453
	β	2.9532	3.6693	3.4985	4.1072	6.7266	9.1348	8.7707	8.2656	6.5201
	a	1.4282	1.0775	0.9442	0.7836	0.5227	0.3485	0.2891	0.2387	0.2259

表 2 暴雨强度抽样各概率分布模式拟合指标

模 式		历 时 (m in)								平 均	
		5	10	15	20	30	45	60	90		120
耿贝尔分布	σ	0.0699	0.0577	0.0737	0.0577	0.0313	0.0309	0.0315	0.0356	0.0414	0.0477
	$f(\%)$	3.5845	3.8220	5.7548	5.1797	3.5203	4.4778	5.5458	8.2971	11.652	5.7594
	g	0.0365	0.0406	0.0580	0.0430	0.0231	0.0247	0.0241	0.0285	0.0306	0.0343
指数分布	σ	0.0565	0.0244	0.0256	0.0263	0.0262	0.0281	0.0214	0.0163	0.0205	0.0272
	$f(\%)$	2.9003	1.6149	1.9998	2.3592	2.9447	4.0700	3.7582	3.7884	5.7773	3.2459
	g	0.0397	0.0176	0.0204	0.0134	0.0175	0.0181	0.0153	0.0133	0.0139	0.0188
韦伯分布	σ	0.0610	0.0377	0.0427	0.0386	0.0237	0.0230	0.0229	0.0230	0.0235	0.0329
	$f(\%)$	3.1298	2.4968	3.3378	3.4624	2.6704	3.3418	4.0342	5.3644	6.6126	3.8278
	g	0.0292	0.0176	0.0274	0.0164	0.0138	0.0193	0.0189	0.0192	0.0121	0.0193
皮尔逊Ⅲ分布	σ	0.04199	0.03428	0.03100	0.02840	0.02332	0.01932	0.01786	0.01517	0.01410	0.02500
	$f(\%)$	2.1546	2.2722	2.4212	2.5474	2.6247	2.8020	3.1397	3.5365	3.9665	2.8394
	g	0.0331	0.0378	0.0421	0.0309	0.0294	0.0334	0.0310	0.0278	0.0182	0.0315

表 3 指数分布 $i-t-T$ 表

重现期 (a)	历 时 (m in)								
	5	10	15	20	30	45	60	90	120
0.25	1.5175	1.1542	0.9601	0.8217	0.6482	0.4913	0.3850	0.2726	0.2109
0.33	1.6404	1.2552	1.0514	0.9053	0.7167	0.5477	0.4373	0.3172	0.2521
0.5	1.8244	1.4063	1.1880	1.0303	0.8191	0.6322	0.5157	0.3838	0.3137
1	2.1312	1.6585	1.4158	1.2388	0.9900	0.7732	0.6464	0.4951	0.4166
2	2.4380	1.9107	1.6436	1.4473	1.1609	0.9141	0.7772	0.6063	0.5195
3	2.6175	2.0582	1.7769	1.5693	1.2608	0.9966	0.8537	0.6714	0.5796
5	2.8437	2.2440	1.9448	1.7230	1.3868	1.1004	0.9500	0.7534	0.6555
10	3.1505	2.4962	2.1727	1.9315	1.5576	1.2414	1.0807	0.8646	0.7583
20	3.4573	2.7483	2.4005	2.1401	1.7285	1.3823	1.2115	0.9758	0.8612
50	3.8629	3.0817	2.7017	2.4158	1.9544	1.5686	1.3843	1.1229	0.9972
100	4.1698	3.3338	2.9295	2.6243	2.1253	1.7096	1.5150	1.2341	1.1000

模式参数求出后,即可确定暴雨强度抽样各概率模式的分布函数,从而利用分布函数求解各个重现期或某一概率对应的暴雨强度.制作反映暴雨强度、历时、重现期这三者关系的 $i-t-T$ 表.表 3 为暴雨强度抽样指数分布的 $i-t-T$ 表.

3 瑞安市暴雨强度公式的构造

3.1 非线性模型参数估计

线性模型的参数求解是根据最小二乘法作估计.假设 $y_i, \hat{y}_i, Q(b_0, b_1, b_2, \dots, b_m)$ 分别表示观测值、回归值和残差平方和, $b_0, b_j (j = 1, 2, 3, \dots, m)$ 是回归方程中的参数.则:

$$Q(b_0, b_1, b_2, \dots, b_m) = \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (21)$$

由于该式是 $b_0, b_1, b_2, \dots, b_m$ 的二次函数,又是非负的,故其最小值总存在.根据微分学的极值定理, $b_0, b_1, b_2, \dots, b_m$ 应是下列联立方程组的解:

$$\begin{cases} \frac{\partial Q}{\partial b_0} = -2 \sum_{i=1}^N (y_i - \hat{y}_i) = 0 \\ \frac{\partial Q}{\partial b_j} = -2 \sum_{i=1}^N (y_i - \hat{y}_i) X_{ij} = 0 \end{cases} \quad (22)$$

若回归方程是非线性,则需要根据给出的几组记录值 $(x_k, y_k) (k = 1, 2, 3, \dots, m)$, 在最小二乘法意义下确定各参数 b_1, b_2, \dots, b_m , 这就要用非线性二乘法来求解,也即确定参数 b_1, b_2, \dots, b_m , 使 $f(x_k, b_1, b_2, \dots, b_m)$ 与 y_k 之间的残差平方和

$$Q(b_1, b_2, \dots, b_m) = \sum_{i=1}^N [y_i - f(x_k, b_1, b_2, \dots, b_m)]^2 \quad (23)$$

为最小.令

$$b_i = b_i^{(0)} + \Delta_i \quad (i = 1, 2, 3, \dots, m) \quad (24)$$

此处 $b_i^{(0)}$ 是 b_i 的一个初始值.

将 $f(x_k, b_1, b_2, \dots, b_m)$ 在点 $b_1^{(0)}, b_2^{(0)}, \dots, b_m^{(0)}$ 展开为泰勒(Taylor)级数,并略去 2 次以上的展开项,得:

$$f(x_k, b_1, b_2, \dots, b_m) \approx f_{k0} + \frac{\partial f_{k0}}{\partial b_1} \Delta_1 + \frac{\partial f_{k0}}{\partial b_2} \Delta_2 + \dots + \frac{\partial f_{k0}}{\partial b_m} \Delta_m \quad (25)$$

其中 $f_{k0} = f(x_k, b_1^{(0)}, b_2^{(0)}, \dots, b_m^{(0)})$

$$\frac{\partial f_{k0}}{\partial b_i} = \left. \frac{\partial f(x, b_1, b_2, \dots, b_m)}{\partial b_i} \right|_{x=x_k}, \quad b_i = b_1^{(0)}, b_2^{(0)}, \dots, b_m^{(0)} \quad (26)$$

令

$$Q = \sum_{i=1}^N [y_k - f(x_k, b_1, b_2, \dots, b_m)]^2 \approx \sum_{i=1}^N [y_k - (f_{k0} + \frac{\partial f_{k0}}{\partial b_1} \Delta_1 + \frac{\partial f_{k0}}{\partial b_2} \Delta_2 + \dots + \frac{\partial f_{k0}}{\partial b_m} \Delta_m)]^2 \quad (27)$$

为了使 Q 在最小二乘法意义下达到极小,对上式 Δ_i 取偏导数,并令其等于零,则有:

终用于构造公式的分布模式. 表 5 和式(32)分别为通过指数分布 $i-t-T$ 表得到的暴雨强度分公式和总公式, 其中总公式的均方根误差为 0.0507(见表 4).

表 4 总公式参数估计值

模 式	参 数				
	A_1	C	b	n	σ^*
指数分布	9.3927	0.5336	6.6612	0.6171	0.0507
韦伯分布	10.1160	0.4168	7.7089	0.6420	0.0747
P-III(权函数法)	9.4274	0.4537	7.5384	0.6196	0.0790
P-III(适线法)	8.6585	0.4522	5.9225	0.6086	0.0673

* σ 为均方根误差.

表 5 瑞安市暴雨强度分公式

重 现 期(a)	分 公 式	均方根误差
0.25	$i = \frac{17.60416}{(10.41968 + t)^{0.89806}}$	0.0106
0.33	$i = \frac{16.00124}{(9.821947 + t)^{0.846846}}$	0.0106
0.5	$i = \frac{14.65082}{(9.23222 + t)^{0.786807}}$	0.0137
1.0	$i = \frac{12.906756}{(7.828816 + t)^{0.707403}}$	0.0106
2.0	$i = \frac{12.30079}{(6.959083 + t)^{0.653405}}$	0.0122
3.0	$i = \frac{12.673768}{(6.814957 + t)^{0.639711}}$	0.0154
5.0	$i = \frac{12.613421}{(6.367485 + t)^{0.613593}}$	0.0200
10.0	$i = \frac{12.962208}{(5.76567 + t)^{0.591829}}$	0.0237
20.0	$i = \frac{13.775758}{(5.498366 + t)^{0.580096}}$	0.0215
50.0	$i = \frac{16.466284}{(6.692008 + t)^{0.588019}}$	0.0304
100.0	$i = \frac{16.2724}{(6.179103 + t)^{0.564611}}$	0.0376

瑞安市暴雨强度总公式:

$$i = \frac{9.392731 + 5.0124 \lg T}{(6.661224 + t)^{0.617077}} \quad (32)$$

4 结 论

暴雨强度公式编制首先是选择适当的概率模式对抽样资料进行拟合并计算出 $i-t-T$ 表; 然后在 $i-t-T$ 表基础上, 利用非线性模型参数估计确定暴雨强度公式中的参数, 即构造各重现期与暴雨强度之间的经验公式. 关于暴雨公式中的参数估计, 常用回归法. 摆试

法、线性最小二乘法等方法。本文在制定非线性模型的参数估计算法中, 为提高参数迭代求解的效率利用了最速下降法, 对计算步骤和算式做了精心设计与编程。在采用差商方法求偏导数时, 利用计算机语言中具有高位精度的扩展型实数来处理, 以使误差控制在很小的范围内, 保证公式参数估计的精度。因此瑞安市暴雨强度公式达到了较高的精度。

参 考 文 献

- 1 中华人民共和国国家标准: 室外排水设计规范, GBJ14- 87, 1988.
- 2 给水排水设计手册(第五册). 北京: 中国建筑工业出版社, 1985, 48~ 87.
- 3 马开玉, 丁裕国, 屠其璞, 等. 气候统计原理与方法. 北京: 气象出版社, 1993, 391~ 419.
- 4 钱焕延. 计算方法. 上海: 上海交通大学出版社, 1988, 297~ 312.

RESEARCH ON PARAMETER ESTIMATE OF PROBABILITY DISTRIBUTION FORMULA OF STORM INTENSITY IN RUIAN

Gu Junqiang Chen Haiyan Xu Jiyun

(Climate Center of Zhejiang Province, Hangzhou 310017)

Abstract

Based on the 33-year (1965-1997) automatically recorded precipitation data of Ruian city, the frequency distribution of storm intensity is discussed by using probability models. The $r-t-T$ table that describes the relationship of storm intensity, rainfall duration and return period is presented. Furthermore, Parameters in the storm intensity formula are estimated by using the optimal solution to solve the super-nonlinear coupled equation.

Key words: Probability model Parameter estimate Storm intensity formula