第13卷 第1期
1990年3月

南京气象学院学报
Journal of Nanjing Institute of Meteorology

Vol.13, No.1
Mar., 1990

N日降水量的随机分布模式

张耀存*

提　要

本文从逐日降水过程干湿日演变的一阶Markov链及日降水量的Γ分布模式出发，在理论上得到任意给定时期N日降水总量及最大日降水量的理论分布函数，同时得到N日总降水量和最大日降水量的数学期望和方差。对所得模式利用上海、北京、广州、沈阳和荆州等5个代表站的实测资料验证，拟合效果较好，由此初步证实所得模式的普适性。因此，可利用该模式估计任意给定时期内N日降水总量和最大日降水量的概率以及相应的降水保证率。

研究一定时期内降水量的总体统计特征对于农业生产的合理布局、水资源的有效开发和利用以及大型水利工程的设计具有重要意义，而从定期降水量的概率分布模式进行研究是一条切实可行的途径。国外学者早在50年代就开始进行这方面的研究，并取得了一些有意义的结论[1-5]。么枕生在其著作中比较系统地讲述了分布函数模式在气候统计分析中所具备的理论，并总结了前人在这方面所做的工作[6]。但国内具体的研究工作还不多，因此有必要进一步开展降水量分布模式理论和实际应用的研究，以便对我国不同时段内的降水过程的总体统计特征有一比较全面的了解。

由于逐日降水资料的记录年限较长，易于获得可靠的资料，因此基于逐日降水过程的统计分布模式建立任意给定时期N日降水量的理论分布模式，用以推算该统计时段内降水过程的统计分布特征，是一项有重要实际意义的工作。本文从逐日降水过程的干湿日演变Markov链模型及假设日降水量服从Γ分布出发，得到任意给定时期N天降水总量及最大日降水量的随机分布模式和N天降水总量及最大日降水量的数学期望及方差。根据所得模式利用5个代表站的实测资料验证，拟合效果较好，初步证实该模式的可靠性。

一、几个有关的概念

考虑由N天构成的某段时期，在这N天中首先定义一个只取0和1两个值的随机变量 η_j，如果N天中第j天是湿日(日降水量大于或等于0.1毫米定义为湿日，否则为干

1988年7月20日收到，1988年10月25日收到修改稿

* 本院88届硕士，现为南京大学大气科学系博士研究生

日)，则以1表示，如果第j天是干日，则以0来表示，即

$$\eta_j=\begin{cases}0 & 若第j天为干日\\ 1 & 若第j天为湿日\end{cases} \tag{1}$$

其中 $j=1, 2, \cdots\cdots, N$。

再以N_n表示N日时期内的雨日天数，则有

$$N_n=\sum_{j=1}^{N}\eta_j \tag{2}$$

显然随机变量N_n可能取的值为0，1，2，……，N。

N日时期内第ν个雨日的降水量用ξ_ν表示，这段时期内的总降水量以S(n)表示，根据(1)、(2)两定义可有

$$S(n)=\sum_{\nu=0}^{N_n}\xi_\nu,\quad \xi_0\equiv 0 \tag{3}$$

由此可得$P\{S(n)=0\}=P\{\eta_1=0,\eta_2=0,\cdots,\eta_n=0\}$以及$P\{S(n)>0\}=1-P\{S(n)=0\}$。

如果以$\chi(n)$表示N日时期内的最大日降水量，即

$$\chi(n)=SUP\xi_\nu \qquad 0\leqslant\nu\leqslant N_n \tag{4}$$

同样有$P\{\chi(n)=0\}=P\{\eta_1=0,\eta_2=0,\cdots\cdots,\eta_n=0\}$以及$P\{\chi(n)>0\}=1-P\{\chi(n)=0\}$。由(3)、(4)两式可知，$S(n)$、$\chi(n)$是离散参数的随机过程，而且有$0\leqslant S(1)\leqslant S(2)\leqslant\cdots\cdots\leqslant S(n)$及$0\leqslant\chi(1)\leqslant\chi(2)\leqslant\cdots\cdots\leqslant\chi(n)$。

由此在一定的假设条件下，可推导出$S(n)$、$\chi(n)$的理论分布函数及其数学期望和方差的表达式。

二、N日降水总量及最大日降水量的理论分布模式

根据以上定义知道，事件序列$\{N_n=0\}$、$\{N_n=1\}$、……、$\{N_n=N\}$代表样本空间的一个有限划分，即对每一个$i\neq j$有$\{N_n=i\}\cap\{N_n=j\}=\theta$，而且

$$\sum_{\nu=0}^{N}P\{N_n=\nu\}=1 \tag{5}$$

式中θ表示空集。下面进一步假设：

(1)ξ_1、ξ_2、……、ξ_n为同分布的独立随机变量，每日降水量的概率分布函数为$H(x)=P\{\xi_\nu\leqslant x\}$。

(2)ξ_1、ξ_2、……、ξ_n对于N_n是独立的。

根据以上假设应用日降水量的分布函数$H(x)$与概率$P\{N_n=k\}$(k为N天中出现的湿日天数，且$k\leqslant N$)，即可确定$\chi(n)$和$S(n)$的理论分布函数及其数学期望和方差[6][7]，这里给出各量的计算表达式。

N日时期内降水总量S(n)的理论分布函数为

$$F_n(x)=P\{S(n)\leqslant x\}$$
$$=P\{N_n=0\}+\sum_{k=1}^{N}P\{X_k\leqslant x\}\cdot P\{N_n=k\} \tag{6}$$

其中$X_k=\xi_1+\xi_2+\cdots+\xi_k$，$k=0,1,2,\cdots,N$，且$X_0\equiv 0$。

N日时期内最大日降水量的理论分布函数为

$$\Phi_n(x)=P\{\chi(n)\leqslant x\}$$
$$=P\{N_n=0\}+\sum_{k=1}^{N}P\{N_n=k\}\cdot\{H(x)\}^k \tag{7}$$

式中的H(x)为日降水量的概率分布函数。

N日时期内降水总量的数学期望和方差的表达式为

$$E\{S(n)\}=\alpha E\{N_n\} \tag{8}$$

$$Var\{S(n)\}=\beta E\{N_n\}+\alpha^2 Var\{N_n\} \tag{9}$$

其中$\alpha=E\{\xi_v\}$，$\beta=Var\{\xi_v\}<\infty$。

N日时期内最大日降水量的数学期望和方差为

$$E\left\{\chi(n)\right\}=\sum_{k=1}^{N}kP\left\{N_n=k\right\}\cdot\left[\int_0^x\left\{H(x)\right\}^{k-1}h(x)dx\right] \tag{10}$$

$$Var\left\{\chi(n)\right\}=\sum_{k=1}^{N}kP\left\{N_n=k\right\}\cdot\left[\int_0^x x^2\left\{H(x)\right\}^{k-1}h(x)dx\right] \tag{11}$$

式中$h(x)=dH(x)/dx$。

三、对概率$P\{X_k\leqslant x\}$与$P\{N_n=k\}$计算公式的改进

为了确定N日时期内的总降水量及最大日降水量的理论分布函数$F_n(x)$、$\Phi_n(x)$，必须计算N天内k个湿日降水量之和的概率分布$P\{X_k\leqslant x\}$及同一时期内出现k个湿日的概率$P\{N_n=k\}$，$k=0,1,\cdots\cdots,N$。Todorovic和Woolhiser(1975)曾假设日降水量服从指数分布，得到了N日时期内总降水量及最大日降水量的理论分布函数模式[7]。作者用指数分布模式对所选代表站拟合日降水量的试验结果表明，指数分布并不能很好拟合日降水量的概率分布，因此文献[7]中得到的模式其通用性必然受到一定的限制。为了得到适用范围较广的N日时期内降水总量及最大日降水量的理论分布模式，本文假设日降水量服从Γ分布，即

$$H(x)=\frac{1}{\beta^\alpha\Gamma(\alpha)}x^{\alpha-1}e^{-x/\beta} \tag{12}$$

式中的 $x>0$，α、β 为形状参数和尺度参数，且都为正值。考虑到 ξ_1、ξ_2、…，ξ_k 相互独立的假设，根据文献[8]，两个湿日降水量和的概率密度表达式为

$$f_{\xi_1+\xi_2}(x)=\int_{-}f_{\xi_1}(x-u)f_{\xi_2}(u)du$$

$$=\int_0 f_{\xi_1}(x-u)f_{\xi_2}(u)du$$

$$=\frac{1}{\Gamma(\alpha)\Gamma(\alpha)\beta^{\alpha+\alpha}}\int_0^x(x-u)^{\alpha-1}e^{-(x-u)/\beta}u^{\alpha-1}e^{-u/\beta}du$$

$$=\frac{e^{-x/\beta}}{\Gamma(\alpha)\Gamma(\alpha)\beta^{\alpha+\alpha}}\int_0^x(x-u)^{\alpha-1}u^{\alpha-1}du$$

对上式进行变换，令 $u=tx$，代入可得

$$f_{\xi_1+\xi_2}(x)=\frac{1}{\Gamma(\alpha+\alpha)\beta^{\alpha+\alpha}}x^{\alpha+\alpha-1}e^{-x/\beta} \tag{13}$$

即两个湿日降水量之和服从参数为 $(\alpha+\alpha,\ \beta)$ 的 Γ 分布。应用数学归纳法，假设 n 个湿日降水量之和服从形状参数为 $n\alpha$、尺度参数为 β 的 Γ 分布

$$f_{\xi_1+\xi_2+\cdots\cdots+\xi_n}(x)=\frac{1}{\Gamma(n\alpha)\beta^{n\alpha}}x^{n\alpha-1}e^{-x/\beta} \tag{14}$$

则 n + 1 个湿日降水量之和的概率密度函数为

$$f_{\xi_1+\xi_2+\cdots\cdots+\xi_{n+1}}(x)=\int_{-\infty}^{\infty}f_{\xi_1+\xi_2+\cdots\cdots+\xi_n}(x-u)\,f_{\xi_{n+1}}(u)\,du$$

$$=\frac{e^{-x/\beta}}{\Gamma(n\alpha)\Gamma(\alpha)\beta^{(n+1)\alpha}}\int_0^x(x-u)^{n\alpha-1}u^{\alpha-1}du$$

作变换 $u=tx$，代入上式，则得

$$f_{\xi_1+\xi_2+\cdots\cdots+\xi_{n+1}}(x)=\frac{1}{\Gamma[(n+1)\alpha]\beta^{(n+1)\alpha}}x^{(n+1)\alpha-1}e^{-x/\beta} \tag{15}$$

即 $\xi_1+\xi_2+\cdots+\xi_{n+1}$ 服从形状参数为 $(n+1)\alpha$、尺度参数为 β 的 Γ 分布。因此可有

$$P\{X_k\leqslant x\}=\frac{1}{\Gamma(k\alpha)\beta^{k\alpha}}\int_0^x u^{k\alpha-1}e^{-u/\beta}du \tag{16}$$

这就是N日时期内 k 个湿日降水量之和概率分布的计算表达式，在实际计算过程中可采用数值积分方法。

Gabriel 和 Neumann 早在1962年就已提出计算N天内出现 k 个湿日概率 $P\{N_n=k\}$ 的方法[9]，Katz(1974)在此基础上作了改进，提出一种递推算法[10]。本文将采用Katz的计算方案[10]。假设逐日降水过程的干湿日演变符合一阶Markov链模型，其状态转移概率为

$$p_{ij}=P\{x_n=j\,|\,x_{n-1}=i\} \tag{17}$$

其中x_n表示N日时期内第n天的干湿状况，i，j＝0，1。

令 $\psi_0(k;\ N)=P\{N_n=k|x_0=0\}$，$\psi_1(k;\ N)=P\{N_n=k|x_0=1\}$，$k=0,1,\cdots,N$。如果湿日的初始概率为 $p=P\{x_0=1\}$，则有

$$P\{N_n=k\}=(1-p)\psi_0(k;\ N)+p\psi_1(k;\ N) \tag{18}$$

式中的 $\psi_0(k;\ N)$、$\psi_1(k;\ N)$可通过以下递推关系式求得

$$\psi_0(k;\ N)=p_{00}\psi_0(k;\ N-1)+p_{01}\psi_1(k-1;\ N-1) \tag{19}$$

$$\psi_1(k;\ N)=p_{10}\psi_0(k;\ N-1)+p_{11}\psi_1(k-1;\ N-1) \tag{20}$$

其中 $k=0,\ 1,\ \cdots,\ N$，并且有初始条件

$$\psi_0(0;\ 0)=\psi_1(0;\ 0)=1,\ \psi_0(N;\ N-1)=\psi_1(-1;\ N-1)=0 \tag{21}$$

同时还可求得N天内的湿日平均天数

$$E\{N_n\}=NQ+(p-Q)\cdot\frac{1-d^N}{1-d}d \tag{22}$$

方差为

$$Var\{N_n\}=Q(1-Q)[N(1+d)/f-2df^2(1-d^N)] \tag{23}$$

式中的$d=p_{11}-p_{00}$，$Q=p_{01}/(1-d)$，$f=1/(1-d)$，p为湿日的初始概率。

四、计　算　结　果

应用上述方法本文对上海、北京、广州、沈阳和荆州等5个代表站分别计算了10天、20天、30天内的总降水量及最大日降水量的理论分布函数值，并且利用实测资料进行验证。模式中参数由各站历史资料应用极大似然法求得。

（1）不同时段内的雨日天数　利用(18)式计算了各站不同时段内雨日出现的频率。图1给出北京、上海两站7月份10天、20天、30天内雨日天数的理论频率分布曲线和由实际资料统计得到的频率分布曲线，干湿日演变的Markov链模型参数见表1。由图1可见，利用(18)式计算得到的雨日频率与实际出现的雨日频率较为一致，同时还可知道北京7月份10天内出现4个雨日的机会较大，而上海出现3天雨日的机会较大。对于20天、30天这两个时段而言，北京、上海两站分别出现9天、13天和6

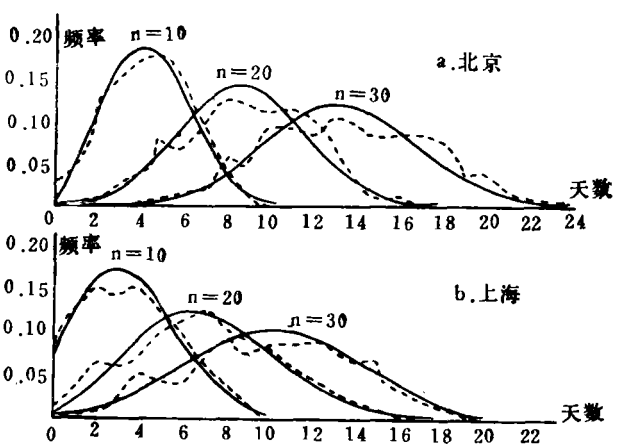

图1　雨日频率分布曲线（7月份）
---理论计算值　——实际观测值

天、10天雨日的可能性较大，这正是雨季气候特征在雨日频率分布上的反映。这些气候统计参数一般不能从气候要素统计表中直接查得，因此探讨它们的间接计算方法对于农业生产部门具有重要的实际意义。

表1 Markov链转移概率值

参数 / 台站	p_{00}	p_{01}	p_{11}	p_{01}
北 京	0.7767	0.2233	0.5923	0.4077
上 海	0.6511	0.3489	0.5594	0.4406

表2 拟合日降水量Γ分布的参数值

参数 / 台站	形状参数	尺度参数
北 京	0.50541	33.04951
上 海	0.53167	24.65584

（2）逐日降水量的概率分布　如前所述，逐日降水量的概率分布可用Γ分布模式拟合，其形状参数和尺度参数用极大似然法求得，表2为北京、上海两站7月份日降水量Γ分布参数的极大似然估计值。图2给出两站7月份日降水量的理论分布曲线和经验分布曲线。由图可见，用Γ分布模式拟合日降水量的理论分布其结果与实际观测资料的经验频率分布比较一致。利用分布的柯尔莫哥洛夫—斯米尔诺夫统计检验法进行拟合适度检验[11]（见表3），由此可以知道，对于α=0.05的置信水平均能通过统计检验，因此用Γ分布模式拟合日降水量的概率分布是可行的。

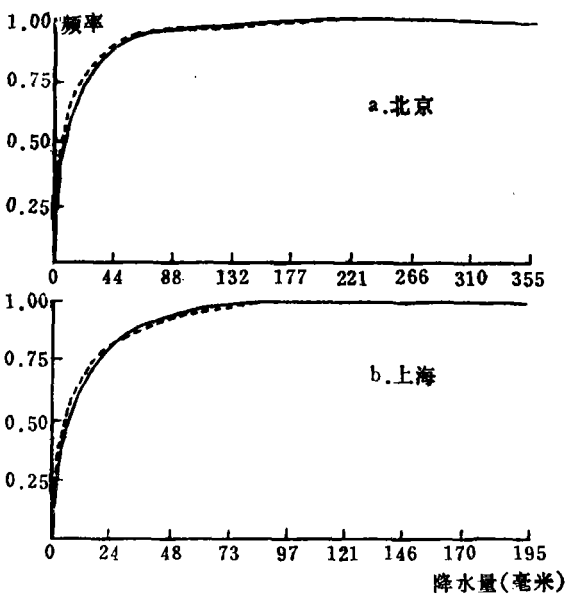

图2 逐日降水量的理论分布(Γ分布)与经验分布曲线(7月份)
——理论曲线　——经验曲线

表3 柯尔莫哥洛夫—斯米尔诺夫适度检验统计量(K—S统计量)

统计量 / 台站	K—S值	α=0.05时的K—S临界值
北 京	1.06014	1.36
上 海	1.08691	

（3）N日时期内总降水量的概率分布　应用（6）式分别对所选各代表站各月计算了10天、20天、30天内总降水量的累积概率分布。图3为北京、上海两站7月份10天、20天、30天内降水总量的累积理论频率分布曲线和累积经验频率分布曲线。由图可见，利用（6）式计算得到的10天、20天、30天内降水总量的概率分布与实测资料统计所得结果比较吻合。统计检验结果表明，各代表站不同月份的计算结果均能通过5%显著性水平的柯尔莫哥洛夫—斯米尔诺夫适度检验，由此可知该模式对位于不同气候区的台站都能得到令人满意的结果，因而其通用性能较好。根据这一模式可计算N天内总降水量超过一定数值的百分位数及降水保证率，从而为农业、水资源开发利用

等提供所需的气候资料服务。

(4)最大日降水量的概率分布　利用(7)式分别计算各站不同月份的最大日降水量的累积概率分布。图 4 给出北京、上海两站 7 月份10天、20天、30天内最大日降水量的理论分布曲线及经验分布曲线。由图可以看到，理论计算与实际观测结果较为吻合。统计检验结果表明，各代表站不同月份的计算结果都能通过 5 %显著性水平的拟合适度检验，因此用该模式计算N日时期内最大日降水量的理论分布是可行的。

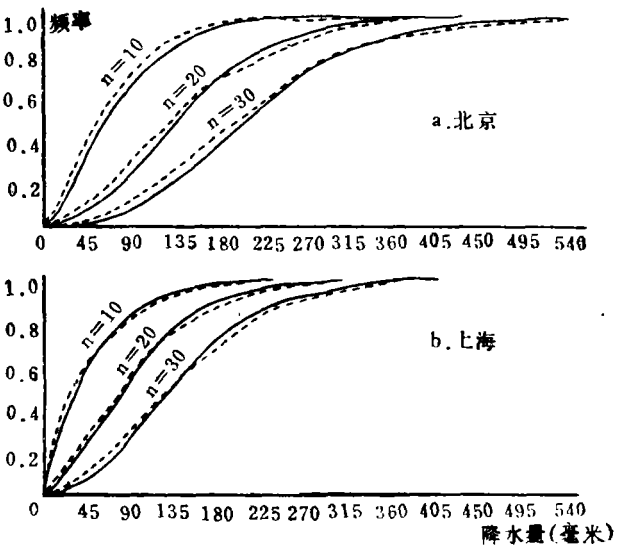

图 3　N日降水总量的理论分布及经验分布曲线（7 月份）
——理论曲线　——经验曲线

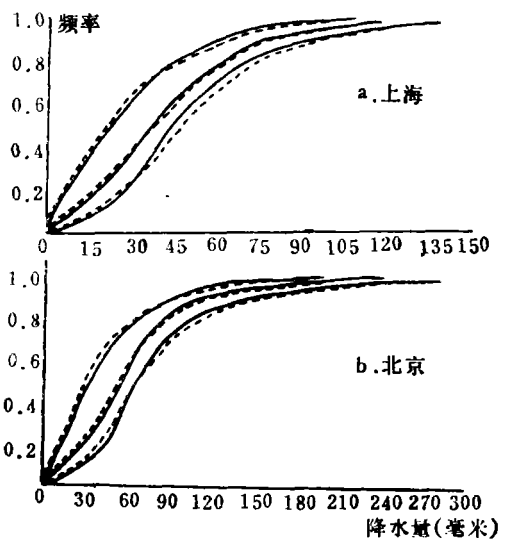

图 4　N天内最大日降水量的理论分布及经验分布曲线（7 月份）
——理论曲线　——经验曲线

五、结 语

本文从逐日降水过程的干湿日演变符合一阶 Markov 链模型及日降水量服从Γ分布的假设出发，从理论上推导得到了任意给定N日时期内降水总量及最大日降水量的随机分布模式，对模式中的几个主要计算方案作了改进，经实测资料验证表明，该模式具有计算结果准确、适用性较广等优点，因此可用以计算N日时期内总降水量及最大日降水量的概率及相应的降水保证率，从而为农业生产、水资源利用及水利工程设计部门提供一种定期降水量概率分布资料的间接计算模式。

本文建立的N日降水量随机分布模式属于气候统计模式的范畴，其计算结果为气候统计指标。在逐日降水过程的干湿日演变 Markov 链模型中，本文规定（参考降水量的常规观测）大于0.1毫米的降水日为湿日，若改变此临界值，可把该模式推广到降水量超过某一

界限的概率分布计算中，这对于实际应用部门具有重要的意义。关于逐日降水过程的干湿日演变符合一阶 Markov 链模型及降水量服从Γ分布假设的合理性问题将另文讨论。

本文曾得到丁裕国副教授热情指导，谨表谢意。

参考文献

[1] Thom, H.C., A note on the gamma distribution, Mon.Wea.Rev., Vol.86, 117—122, 1958.

[2] Caskey, J.E., A Markov chain model for the probability of precipitation occurrence in intervals of various length, Mon.Wea.Rev., Vol.91, 298-301, 1963.

[3] Feyerherm, A. M., Bark, L.D., Statistical methods for persistent precipitation patterns, J.Appl.Met., Vol. 4, 320—328, 1965.

[4] Ison, N. T., Feyerherm, A.M.and Bark, L.D., Wet period precipitation and the gamma distribution, J.Appl.Met., Vol.10, 658—665, 1971.

[5] Mooley, D.A.,Gamma probability model for Asian Summer monsoon monthly rainfall, Mon.Wea.Rev., Vol.101, 160—176, 1973.

[6] 么枕生，气候统计学基础，科学出版社，1984.

[7] Todorovic,P.,Woolhiser,D.A., A stochastic model of n—day precipitation,J.Appl. Met., Vol.14, 17—24, 1975.

[8] 周概容，概率论与数理统计，高等教育出版社，1984.

[9] Gabriel, K.R., Neumann, J., A Markov chain model for daily rainfall occurrence at Tel Aviv, Quart.J.R.Met.Soc., Vol.88, 90—95, 1962.

[10] Katz, R.W., Computing probabilities associated with the Markov chain model for precipitation, J.Appl.Met., Vol.13, 953—954, 1974.

[11] 屠其璞、丁裕国等，气象应用概率统计学，气象出版社，1984.

A STOCHASTIC DISTRIBUTION MODEL OF N—DAY PRECIPITATION AMOUNT

Zhang Yaocun

ABSTRACT

Theoretical expressions are derived for the distribution function of as well as the mathematical expectation and variance of the total amount of precipitation and the daily maximum precipitation occurring in an N-day period based on the one-order Markov Chain and the Gamma probability model of daily rainfall. By using the data from five representative stations over China, we find that the model has a good fitting effect and can thus be used to estimate the probability of the total amount of precipitation and the daily maximum precipitation in any N-day period.