第 31 卷　第 2 期(424 ~ 431)　中 国 地 震　Vol. 31　No. 2
2015 年 6 月　EARTHQUAKE RESEARCH IN CHINA　Jun. 2015

孙丽娜、齐玉妍、金学申，2015，华北地区强震的等待时间研究，中国地震，**31**(2)，424 ~ 431。

华北地区强震的等待时间研究

孙丽娜　齐玉妍　金学申

河北省地震局，石家庄市槐中路 262 号　050021

摘要　从概率论出发，推导出从上一次地震发生时间到下一次地震发生时间(等待时间)的数学表达式，并给出了威布尔分布、指数分布、均匀分布的等待时间表达式。在此基础上，以河北省及其邻近地区为例，用 3 个分布函数分别计算了各地区发生 5.0 级以上地震的等待时间发震概率及回溯性检验，并利用 AIC 准则对何种分布函数更适合华北地区 5.0 级以上强震危险性分析进行了讨论。

关键词：　强震　概率　分布　等待时间

［文章编号］ 1001-4683(2015)02-0424-08　　［中图分类号］ P315　　［文献标识码］ A

0　引言

定量化的地震预测模型及其检验是推进对地震可预测属性的科学认识、提高区域地震危险性分析水平的关键，国际上正在开展的“地震科学可预测性合作研究计划”(Collaboratory for the Study of Earthquake Predictability，CSEP)对此开展了大量工作(Jordan，2006；蒋长胜等，2009)。

强震的孕育、发生是极其复杂的物理过程，离上次发生强震的时间越长，离下一次地震的时间不一定也越长。Garavaglia 等(2007)的工作指出，如果考虑强震的发生只取决于从上一强震之后的若干时间发生，那么在某区域的地震预测是可能的。

假设某地区在某时间发生了一次强震，对于如何判断下一次地震的发生时间，许多研究者从不同的角度对不同环节提出了各种有益的探索。然而，就地震复发模型而言至今仍主要沿用泊松过程的假定。当去除余震后，小地震活动符合泊松分布(Gardner et al，1974)，而在一个较大范围里的中强地震活动在时间轴上却遵从成丛分布(Knopoff et al，1996)。按照这个假设，周期性的地震活动遵从有限的成丛性，可以精确地预测。一般来说，如果有时间上的聚集性，且在已有地震时间分布的前提下，那么就可以计算未来一定时间间隔的发震概率。为了定量描述这一性质，我们把在给定区域自上一次强震以来的持续时间与在给定时间间隔范围内的强震发生的条件概率联系起来，计算强震在给定时间的发生概率。

本文从概率论的原理出发，推导出计算强震等待时间的公式，然后依据各分布模型，得

［收稿日期］ 2014-01-02；［修定日期］ 2014-07-18

［项目类别］ 河北省地震科技星火计划重点课题资助

［作者简介］ 孙丽娜，女，1982 年生，硕士，主要从事地震中长期预测等研究。E-mail：sunlina20082008@126.com

到等待时间的计算公式,在此基础上,计算了河北省及邻近地区的5.0级以上强震等待时间的概率。

1 下一次地震的等待时间

假定 $P(t)$ 是地震时间间隔的概率密度,而 t 是自上一次地震以来的时间,t' 是从现在到下一次地震发生的时间间隔,即等待时间(Sornette et al,1997)。现在的问题是要求得 t' 的概率密度函数 $P(t')$。据贝叶斯条件概率公式

$$P(A/B) = \frac{P(A)}{P(B)} \tag{1}$$

我们假定 $P(A) = P(t + t')$,这就是从现在开始到下一次地震发生时的时间间隔为 t' 的概率。同时假定 $P(B) = \int_t^\infty P(s)\,\mathrm{d}s$,也就是从上一次地震发生到现在没有地震发生的概率。这样,归一化的条件概率公式为

$$P(t') = \frac{P(t+t')}{\int_t^\infty P(s)\,\mathrm{d}s} \tag{2}$$

由(2)式得平均等待时间为

$$\langle t' \rangle = \frac{\int_0^\infty t' P(t+t')\,\mathrm{d}t'}{\int_t^\infty P(u)\,\mathrm{d}u} \tag{3}$$

通过简单的变量变换 $u = t + t'$,可以得到

$$\langle t' \rangle = \frac{\int_t^\infty (u-t) P(u)\,\mathrm{d}u}{\int_t^\infty P(u)\,\mathrm{d}u} \tag{4}$$

由分部积分,得到

$$\langle t' \rangle = \frac{\int_t^\infty \mathrm{d}s \int_s^\infty P(u)\,\mathrm{d}u}{\int_t^\infty P(u)\,\mathrm{d}u} \tag{5}$$

为简单起见,把(5)式写成

$$\langle t' \rangle = -\frac{f(t)}{f'(t)} \tag{6}$$

$f(t)$ 是 $p(u)$ 的二重积分,因此有

$$f(t)f''(t) - [f'(t)]^2 > 0 \tag{7}$$

则

$$\frac{\mathrm{d}\langle t' \rangle}{\mathrm{d}t} > 0 \tag{8}$$

同样,设 $f(t) = \mathrm{e}^{-g(t)}$,如果 $g''(t) < 0$,则

$$\frac{\mathrm{d}\langle t' \rangle}{\mathrm{d}t} = -\frac{g''(t)}{[g'(t)]^2} > 0 \tag{9}$$

如果 $P(t)$ 在 $t=0$ 时有限，我们也能发现(5)式的变换形式，这对等待时间短的情况十分有用。因为对短的等待时间，有 $\lim\limits_{t\to 0^+}\frac{\mathrm{d}\langle t'\rangle}{\mathrm{d}t} = P(0)\Delta - 1$，其中，$\Delta = \int_0^{\infty}\mathrm{d}s\int_s^{-\infty}P(u)\,\mathrm{d}u = \langle t\rangle$。积分结果可以直接据(5)式得到: $\langle t\rangle$ 是两地震间的平均重复时间，设 $\tau = \frac{1}{P(0)}$，其中 τ 是等待时间的估计值，它是根据上一次地震以来的时间估计得到的。我们把它看作为 t' 的瞬时估计。这样 $\frac{\mathrm{d}\langle t'\rangle}{\mathrm{d}t}$ 就可以写成 $\lim\limits_{t\to 0^+}\frac{\mathrm{d}\langle t'\rangle}{\mathrm{d}t} = \frac{\langle t\rangle}{\tau} = 1$。

如果瞬时估计 τ 小于平均的等待时间，那么，到下一次地震的时间随着离逝时间的增加而增加，这表明平均等待时间反应的是整个时间中各部分的贡献。而如果 $\langle t\rangle$ 的值大于 τ，则表明这是在非零时刻大于 τ 的那部分的贡献。在这种情况下，$\lim\limits_{t\to 0^+}\frac{\mathrm{d}\langle t'\rangle}{\mathrm{d}t} > 0$。如果 $\langle t\rangle$ 小于 τ，离下一次地震的时间随着离逝时间的增加而减少。在特殊情况下，如果 $P(0) = 0$，则 $\lim\limits_{t\to 0^+}\frac{\mathrm{d}\langle t'\rangle}{\mathrm{d}t} = -1$，离下一次地震的时间随着离逝时间的增加而减少。如果 $P(0) = \infty$，则 $\lim\limits_{t\to 0^+}\frac{\mathrm{d}\langle t'\rangle}{\mathrm{d}t} = \infty$，即离下一次地震的时间随着离逝时间的增加而增加。

2　各种分布的概率密度函数

假设上次地震起至现在的时间为 t，其概率密度函数为 $p(t)$，现在起至下次发生地震的间隔时间为 t' 的概率密度函数为 $P(t')$；现在起至下次发生地震间隔时间在 u 以内的概率为 $Pr(u)$。则由(2)式可得

$$P(t') = \frac{P(t+t')}{\int_t^{\infty}P(s)\,\mathrm{d}s} \tag{10}$$

$$Pr(u) = \int_0^{u}P(t') \tag{11}$$

根据这些分布函数的定义，结合(2)式和(11)式用分布函数来计算下一次强震的等待时间概率。

2.1　指数分布

我们首先考虑指数分布，因为它与泊松分布相似，指数分布的等待时间的表达式为

$$P(t) = \frac{\mathrm{e}^{-\frac{t}{t_0}}}{t_0},\quad (t \geqslant 0) \tag{12}$$

其中 t_0 是平均间隔时间。由(2)式，可得

$$P(t') = \frac{\mathrm{e}^{-\frac{t'}{t_0}}}{t_0} \tag{13}$$

很显然，下一次地震的发生时间不取决于离逝时间，不管 t 的值有多大，从现在到下一次地震的平均时间都为 t_0。则

$$Pr(u) = 1 - \mathrm{e}^{-\frac{u}{t_0}} \tag{14}$$

2.2　均匀分布

均匀分布的表达式是

$$P(t)=\frac{1}{2t_0},\quad (0\leqslant t\leqslant 2t_0) \tag{15}$$

其中 $2t_0$ 是地震间最大的时间间隔，t_0 是其平均值。由(2)式，有

$$P(t')=\frac{1}{2t_0-t},\quad (0\leqslant t'\leqslant 2t_0-t) \tag{16}$$

$P(t')$ 与 t' 无关，也就是说，它本身是一个均匀分布，但其值取决于 t，在未来一直到 t_0 的时间内，发生一次地震的概率随着离逝时间的增加而增加，而当离逝时间超过 $2t_0$ 时，等待时间就变得不确定了。这实际上也就告诉我们，下一次地震的发生时间在 $2t_0$ 内的可能性最大，而这从公式 $\langle t'\rangle=\frac{1}{2(2t_0-t)}$ 也可以看出。如果 $\frac{\mathrm{d}\langle t'\rangle}{\mathrm{d}t}$ 是负的，这表明等待时间随着离逝时间的减少而减少。则

$$Pr(u)=\frac{u}{2t_0-t} \tag{17}$$

2.3　威布尔分布

威布尔分布的表达式为

$$P(t)=mt_0^{-m_t}t^{m-1}\mathrm{e}^{-(t/t_0)^m},\quad (0<t<\infty,m>0) \tag{18}$$

$$P(t')=mt_0^{-m_t}(t+t')^{m-1}\mathrm{e}^{-\frac{(t+t')^m-t^m}{t_0^m}} \tag{19}$$

$$Pr(u)=1-\mathrm{e}^{-\frac{(t+u)^m-t^m}{t_0^m}} \tag{20}$$

它最有可能的值是 $(m-1)^{1/m}t_0$，平均值为 $\alpha(m)t_0$，其中 $\alpha(m)=\int_0^\infty \mathrm{e}^{-t^m}$。当 $m=1$ 时，威布尔分布有一个特殊情况就是上面所讲的指数分布(梁之舜,1988)。

3　t_0 的确定

在上述计算中，我们不知道 t_0，所以要据有限的地震时间间隔来估计。假定我们有 $(n-1)$ 个地震时间间隔，并忽略其不确定性(而这对历史地震来说是可能的)。假定最后一次地震的离逝时间为 t，那么，在泊松分布的情况下，由最大似然法给出的 t_0 的估值是

$$t_0=\frac{1}{n}\left(t+\sum_{j=1}^{n-1}t_j\right) \tag{21}$$

t 为最后一次地震距今的时间间隔，t_j 为 n 个地震间的时间间隔。

这样，在泊松分布的情况下，得到了随离逝时间的增加而增加的下一次地震复发的平均时间间隔。(21)式可以被推广用于上述其他分布(Sornette et al,1997)。

4　强震的等待时间预测方法在河北及周边地区的应用及回溯性检验

根据黄玮琼(1994)对我国大陆地震资料的完整性研究成果，按华北地区(除海域与边远地区外)的记载能力，对公元 1303～1500 年有记载的 6 级以上的地震和公元 1500 年后有

记载的 5 级以上的地震，用统计分析方法进行检验，得出华北地区（除黄海及边远地区外）$M \geq 4\frac{3}{4}$ 地震自 1484 年之后基本完整。因此，本文将公元 1500 年以来的地震目录（去除余震）用于等待时间的数学模型计算。

本文将预测研究区划分成：河北南部、晋冀蒙、京津、唐山 4 个区域（表 1），用公元 1500 ~1970 年的 5 级以上的地震目录计算了这 4 个地区从 1970 ~ 2070 年这 100 年内各时段发生 5 级及以上地震的概率（图 1）。

表 1 研究区域的划分

河北南部	晋冀蒙	京津	唐山
113° ~116°E	113° ~116°E	115.5° ~117.5°E	117.5° ~119.5°E
34° ~38.5°N	38.5° ~42°N	38.5° ~41.5°N	39° ~41°N

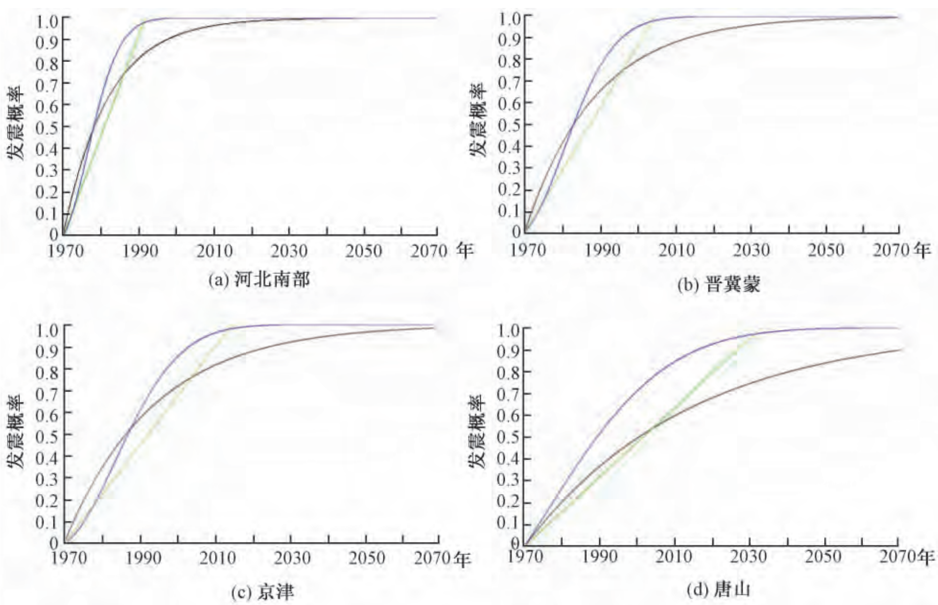

图 1 研究区域 5 级以上地震在 1970 ~2070 年的发震概率

——指数；——均匀；——威布尔

本文设发震概率为 P，当 $P \geq 0.6$ 时，认为将会发生 5 级及以上的地震，则 $P = 0.6$ 是发生地震的最低阈值，因此对 $P = 0.6$、0.7、0.8 这 3 个分布计算的发生 5 级及以上地震的概率值所对应的时间，进行回溯性检验。结果如表 2 ~5 所示。

据表 2 ~5 可以看出，河北南部和晋冀蒙地区检验率比较高，计算出来的发震时间和实际发震时间比较接近甚至一致。京津地区和唐山地区检验率略低一些，分析京津和唐山地区的计算过程发现，这两个地区在 1500 ~ 1970 年间 5 级及以上地震的时间间隔比较大，这可能是造成检验率低的关键因素。

表 2　河北南部地区回溯性检验结果的对比

河北南部	$P=0.6$ 对应的时间	$P=0.7$ 对应的时间	$P=0.8$ 对应的时间
指数分布	1981 年	1984 年	1989 年
威布尔分布	1979 年	1981 年	1983 年
均匀分布	1984 年	1985 年	1988 年
时间段	1970 ~ 1990 年	1980 ~ 1990 年	1980 ~ 1990 年
实际地震发生时间	1981、1983 年	1981、1983 年	1981、1983 年

表 3　晋冀蒙地区回溯性检验的对比结果

晋冀蒙	$P=0.6$ 对应的时间	$P=0.7$ 对应的时间	$P=0.8$ 对应的时间
指数分布	1988 年	1993 年	2000 年
威布尔分布	1986 年	1988 年	1992 年
均匀分布	1991 年	1995 年	1998
时间段	1980 ~ 2000 年	1980 ~ 2000 年	1990 ~ 2000 年
实际地震发生时间	1981、1989、1991、1998、1999 年	1981、1989、1991、1998、1999 年	1991、1998、1999 年

表 4　京津地区回溯性检验结果的对比

京津	$P=0.6$ 对应的时间	$P=0.7$ 对应的时间	$P=0.8$ 对应的时间
指数分布	1992 年	1999 年	2008 年
威布尔分布	1990 年	1994 年	1998 年
均匀分布	1997 年	2002 年	2006 年
时间段	1990 ~ 2000 年	1990 ~ 2010 年	1990 ~ 2010 年
实际地震发生时间	无	2006 年	2006 年

表 5　唐山地区回溯性检验结果的对比

唐山	$P=0.6$ 对应的时间	$P=0.7$ 对应的时间	$P=0.8$ 对应的时间
指数分布	2010 年	2023 年	2040 年
威布尔分布	1994 年	1999 年	2007 年
均匀分布	2008 年	2014 年	2020 年
时间段	1990 ~ 2010 年	1990 ~ 2030 年	2000 ~ 2040 年
实际地震发生时间	1991 年	1991 年	待定

5　AIC 判别准则

统计模型的优劣可以用(Akaike,1997)提出的一种 AIC 准则来判别,此准则实质上是一种最大熵估计。其方法是:假定 $\ln L$ 是所得到的模型的对数最大似然值,则

$$\mathrm{AIC} = -2\ln L + 2K \tag{22}$$

其中,L 为统计分析中的最大似然数,K 为参量数目。对于不同模型计算出的 AIC 值,AIC 值最小的模型作为最优来选择。两个模型的 AIC 值相差越大,模型的优劣差别就越明显,反之差别不大(朱守彪等,2002)。一

表 6　研究区三个分布函数模型的 AIC 值

分布函数	河北南部	晋冀蒙	京津	唐山
指数分布	4.23	4.34	4.52	4.68
均匀分布	10.61	11.67	13.28	13.22
威布尔分布	13.85	6.04	6.14	6.13

般来说，AIC 值相差 2 以上时，这两个模型就可以比较其优劣（孙若昧等，1999）。

本文用 AIC 准则对研究区的 3 个分布函数模型进行判别，计算结果如表 6。比较表 6 的 AIC 值，发现在这 4 个研究区内，3 个分布函数模型中均是指数分布模型的 AIC 值最小，且这 3 个分布函数模型的 AIC 值都相差 2 以上，AIC 准则适用于判断这 3 个分布函数模型的优劣。结合回溯性检验的结果，本文认为指数分布模型比较适合于这 4 个研究区域内强震等待时间的计算。从图 1 的(a)、(b)、(c)、(d)均可以看出在 2013 年以后的发震概率比较高，这为地震的中长期预测提供了参考。

6 讨论和结论

研究发现，如果将本文的方法外推用于渐近估计非常大时间间隔的分布是危险的，而受计算中各个分布函数的特性和参数值的影响，该方法适用于通过短时间间隔的数据统计推断出长时间尺度的强震等待时间的发震概率。

通过实例分析可以发现，重复时间的统计估计对所假定的分布是非常灵敏的，不同分布的等待时间发震概率有明显差别，这可能与各地区历史地震时间分布有关。且本文在用 AIC 准则判别统计模型的优劣时发现，研究所用的 3 个分布函数中，指数分布模型比较适合用于计算强震等待时间。

从计算结果和实例检验结果来看，强震等待时间的计算方法，对于中长期地震活动具有一定的预测能力，在今后的中长期地震预测中有待于进一步研究。

参考文献

黄玮琼，1994，中国大陆地震资料完整性研究之一——分区地震资料基本完整的起始年分布图，地震学报，**16**(4)，273 ~280。

蒋长胜、赵祎喆，2009，地震可预测性研究的 CSEP 计划及其启示，地震地磁观测与研究，**30**(5)，34 ~40。

梁之舜，1988，概率论及数理统计，110 ~113，北京:北京大学出版社。

孙若昧、石耀霖、刘杰，1999，华北地震和波速结构关系分类表格的数据 AIC 分析，地震学报，**21**(1)，10 ~16。

朱守彪、石耀霖，2002，应力释放模型的改进及其在研究台湾地区地震预测问题中的应用，地震学报，**24**(2)，162 ~168。

Akaike H，1977，On entropy maximization principle，In: Krishnaian P R，Applications of Statistics，Amsterdam: North Holland，27 ~41.

Garavaglia E，Guagenti E，Pavani R，et al，2007，Renewal models for earthquake predictability，Journal of Seismology，10.1007/s10950 -008 -9147 -6.

Gardner J K，Knopoff L，1974，Is the sequence of earthquakes in Southern California，with aftershocks removed，Poissonian? Bull Seism Soc Am，64，1363 ~1367.

Jordan T H. 2006，Earthquake predictability，brick by brick，Seism Res Lett，**77**(1)，3 ~6.

Knopoff L，Levshina T，Keilis-Borok V I，et al，1996，Increased long-range intermediate-magnitude earthquake activity prior to strong earthquakes in California，J Geophys Res，**101**，5779 ~5796.

Sornette D，knopoff L，1997，The paradox of the expected time until the next earthquake，Bull Seism Soc Am，**87**(4)，789 ~798.

Research on waiting time of the strong earthquakes in North China

Sun Lina *Qi Yuyan* *Jin Xueshen*

Earthquake Administration of Hebei Province, Shijiazhuang 050021, China

Abstract The paper derives a mathematical expression on the basis of the probability theory for the time from one earthquake to the next earthquake (waiting time), and gives the waiting time expressions in terms of Weibull distribution, exponential distribution, and uniform distribution. We take Hebei Province and its adjacent area as an example, and the three distribution functions are calculated for the waiting time probability of earthquakes with $M \geqslant 5.0$ in the regions in future and the kind of distribution function is more suitable for the North China area. the $M \geqslant 5.0$ earthquake hazard analysis are discussed using AIC criteria.

Key words: **Strong earthquake Probability Distribution Waiting time**