中国精算师资格考试辅导用书

生存模型的构造理论

邹公明　编著

上海财经大学出版社

中国精算师资格考试辅导用书

生存模型的构造理论

邹公明　编著

上海财经大学出版社

图书在版编目(CIP)数据

生存模型的构造理论/邹公明编著. —上海:上海财经大学出版社,2005.8
(中国精算师资格考试辅导用书)
ISBN 7-81098-439-X/F·395

Ⅰ.生… Ⅱ.邹… Ⅲ.人寿保险-精算学-资格考核-自学参考资料
Ⅳ.F840.62

中国版本图书馆CIP数据核字(2005)第084141号

□ 特约编辑 刘 兵
□ 责任编辑 徐 超
□ 封面设计 周卫民

SHENGCUN MOXING DE GOUZAO LILUN
生存模型的构造理论
邹公明 编著

上海财经大学出版社出版发行
(上海市武东路321号乙 邮编200434)
网 址:http://www.sufep.com
电子邮箱:webmaster@sufep.com
全国新华书店经销
苏州望电印刷有限公司印刷装订
2005年8月第1版 2005年8月第1次印刷

850mm×1168mm 1/32 14.5印张 363千字
印数:0 001—3 000 定价:33.00元

中国精算师资格考试辅导用书

编委会

序　言

选择中国精算师资格考试就像是选择了一条充满荆棘的黄金之路，如何排除行进路上的一个个荆棘，是一项具有一定技术的事情。许多考生认为在具有一定数学基础的情况之下，只要勤奋就足够了，事实上这是错误的观点。不少勤奋的考生可以说是饱尝了考试精算的苦痛，另有一部分考生也可能顺利地拿到了精算师证书，可是他们并没有感受到精算的深刻内容，当然更没有感受到他们期望的精算赋予给他们的神奇本领。该停下来反省一下了，忙碌的学子们。

本书丰富的习题试图给学习者暗示了以下排除荆棘的技术，那就是博览群书以弥补精算实务的欠缺以及某一具体精算技术应用的广泛性提示。精算是要在实务中做的，可是芸芸考生之中有几个有机会接触到精算实务呢？而我们国家的精算实务也不过只是个开端而已。博览中外精算专家的著作、文章有助于解决这一学习障碍，研读著名公司编制的财务精算软件说明书，也能窥到一些先进的精算分析方法。本系列丛书中来源广泛的习题力求做到各种知识在精算学科中的应用，并且有些习题甚至就可扩充为一个险种，为实务中的产品开发者提供灵感。因此本书百科全书般的习题可以引领考生步入缤纷多彩的精算实务世界。另外就事论事是学不好精算学的，只会是一叶障目，不见树林，更难见到生态多样的森林。譬如生命表的构造理论中的人口数学并不仅仅是构造生命表之用，其中的人口规划或人口模型应用于保险微观规划可谓是巧夺天工。然而更为深远的用处是该系列模型可应用于保险险种乃至宏观保险经济的规划及预测。即便是生命表的构造理

论也不只是仅用于构造生命表之用，这种方法也可应用于某些非寿险领域，还可应用于疾病保险的开发等等。这系列丛书会给读者暗示某一具体精算技术的广泛性应用，而不是一叶障目的就事论事。这样学习精算的“哥伦布”们不仅发现了美洲大陆，而且缔造了一个伟大的国家——美国，效用何其大也。

也就是说，如果读者得到了本书的真传，在学习精算的过程中，读者将会在闪着金光的但是充满荆棘的路途中感受着艰辛带来的乐趣，之后的若干年，读者会充分享受精算给你带来的收益和快乐。在快乐之余，这套丛书中由于编者的水平局限及时间仓促带来的不足之处，希望读者能够谅解，并且记住曾经有那么几个作者不失时机地给读者提供的精算教育。

邹公明

2005 年 7 月 8 日

前　言

《生存模型的构造理论》是专为参加中国精算师资格考试的考生而写的，本书涉及到的理论及方法是寿险精算的基础。不仅如此，本书的估计理论对日常管理事务的科学判断也是有益的，特别对医学研究及医院管理人员尤其有意义。再者其中的人口理论中众多的例子，也将会使从事养老金规划、团体保险及保险宏观研究人员受到启发，从而开辟新的研究或工作思路。其实，这部分理论还会使营销战略制订人员及生物种群研究者产生新的灵感，因此，此书的读者将不会拘于致力于中国寿险精算师考试的人们，从事非寿险工业产品损坏险的设计人员及对工业产品可靠性有研究的人应该会有同感。

学习本书并不需要太多高深的理论，但读者至少要有一般初等的数理统计基础；对于模型修匀方面知识的学习，读者若有时间序列及数值计算的修养，则会更显得心应手一些。至于马尔可夫随机过程方面的知识，在学习人口规划理论时会有用处。本科目是考生公认的中国精算师资格考试最难的科目之一，读完本书，这种感觉就会烟消云散，并且会感受到这本书给予读者的远不只是构造生存模型之用，因为她会带你徜徉保险业深不见底的知识海

洋，帮你洗去急功近利的浮躁，你的精算人生也会因此而更加稳健和绚丽多彩。

由于作者水平有限，书中错误疏漏之处在所难免，望读者指正。另外在编写本书的过程中，作者得到了多位教授及精算专家的大力帮助和指点，在此表示衷心的感谢。

邹公明

2005 年 6 月

目　录

序言 ………………………………………………………………（1）

前言 ………………………………………………………………（1）

第一章　生存模型及其性质 ………………………………（1）
第二章　生命表 ……………………………………………（19）
第三章　完整样本数据情况下表格生存模型的估计 ………（49）
第四章　非完整样本数据情况下表格生存模型的估计 ……（66）
第五章　参数生存模型的估计……………………………（111）
第六章　大样本数据情况下年龄的处理及暴露数的计算
……………………………………………………………（126）
第七章　人口数学…………………………………………（156）
第八章　修匀数学…………………………………………（226）
全真模拟试题(一)…………………………………………（310）
全真模拟试题(二)…………………………………………（321）
全真模拟试题(三)…………………………………………（331）
全真模拟试题(一)解答……………………………………（343）
全真模拟试题(二)解答……………………………………（366）

全真模拟试题(三)解答…………………………………… (385)
2001 年度中国精算师资格考试试题及解答 ………………… (406)
附录一 周江雄等编著《生命表的构造理论》勘误表……… (445)
附录二 复习顺序图…………………………………………… (449)

后记……………………………………………………………… (450)

第一章　生存模型及其性质

一、知识要点

1. 生存模型

$$S(t;x_1,x_2,\cdots,x_n)=P(T(x_1,x_2,\cdots,x_n)>t)$$

其中 t 是所考察对象的未来寿命；$x_1,x_2,\cdots x_n$ 是对未来寿命有影响的伴随变量。

当 $n=1$，并且 $x_1=0$ 时，上述模型变为：

$$S(t,0)=P(T(0)>t)=P(T>t)$$

此模型可以作为零岁的人的生存模型。

2. 危险率函数

$$\lambda(t)=\mu_t=\frac{f(t)}{S(t)}=\frac{-S'(t)}{S(t)} \tag{1}$$

其中 $f(t)$ 是与 $S(t)$ 相对应的概率密度函数。

由(1)式可知：

$$S(t)=\mathrm{e}^{-\int_0^t \mu_x \mathrm{d}x}$$

对于选择年龄为 x 岁的生存模型 $S(t;x)$ 的危险率函数

$$\mu_{[x]+t}=\frac{-\dfrac{\partial(S(t,x))}{\partial t}}{S(t,x)}=-\frac{\partial(\ln(S(t,x)))}{\partial t}$$

3. 参数生存模型

①均匀分布

②指数分布

③Gompertz 分布

$$\lambda(x)=\mu_x=BC^x, x\geqslant 0, B>0, C>1$$

④Makeham 分布

$$\lambda(x)=A+BC^x, x\geqslant 0, B>0, C>1, A>-B$$

⑤Weibull 分布

$$\lambda(x)=k\cdot x^n, x\geqslant 0, k>0, n>0$$

4. 截尾分布及一些数字特征

设 X 表示"0"岁人的未来余命随机变量，有：

$$\begin{aligned}{}_np_x&=S(x+n\mid X>x)\\&=P(X>x+n\mid X>x)\\&=\frac{S(x+n)}{S(x)}\end{aligned}$$

$$\begin{aligned}{}_nq_x&=F(x+n\mid X>x)\\&=P(X\leqslant x+n\mid X>x)\\&=\frac{S(x)-S(x+n)}{S(x)}\end{aligned}$$

$$\begin{aligned}f(y\mid X>x)&=\frac{\mathrm{d}F(y\mid X>x)}{\mathrm{d}y}\\&=\frac{\mathrm{d}\left(\frac{S(x)-S(y)}{S(x)}\right)}{\mathrm{d}y}\\&=\frac{f(y)}{S(x)}\end{aligned}$$

$$\lambda(y\mid X>x)=\frac{f(y\mid X>x)}{S(y\mid X>x)}=\lambda(y)$$

$$\begin{aligned}S(x\mid y<X\leqslant z)&=P(X>x\mid y<X\leqslant z)\\&=P(x<X\leqslant z\mid y<X\leqslant z)\\&=\frac{S(x)-S(z)}{S(y)-S(z)}\end{aligned}$$

$$\begin{aligned}F(x\mid y<X\leqslant z)&=1-S(x\mid y<X\leqslant z)\\&=\frac{F(x)-F(y)}{F(z)-F(y)}\end{aligned}$$

$$f(x|y<X\leqslant z)=\frac{\mathrm{d}F(x|y<X\leqslant z)}{\mathrm{d}x}$$

$$=\frac{f(x)}{S(y)-S(z)}$$

$$\lambda(x|y<X\leqslant z)=\frac{\lambda(x)\cdot S(x)}{S(x)-S(z)}$$

$$E(X\mid y<X\leqslant z)=\int_y^z x\cdot f(x\mid y<X\leqslant z)\mathrm{d}x$$

$$E(X\mid y<X)=\int_y^{+\infty} xf(x\mid X>y)\mathrm{d}x$$

y 岁的人的期望余命：

$$\mathring{e}_y=E(X\mid X>y)-y$$

$$=\int_y^{+\infty}(x-y)f(x\mid X>y)\mathrm{d}x$$

$$=\int_0^{+\infty}xf(x+y\mid X>y)\mathrm{d}x$$

$$E(X^2\mid X>y)=\int_y^{+\infty}x^2f(x\mid X>y)\mathrm{d}x$$

设 $T(y)$ 是 y 岁的人的余命随机变量，则 $T(y)=X-y\mid X>y$。

5. 中心死亡率

$$m_x=\frac{\int_0^1 {}_tp_x\mu_{x+t}\mathrm{d}t}{\int_0^1 {}_tp_x\mathrm{d}t}$$

$$=\frac{\int_x^{x+1}S(y)\lambda(y)\mathrm{d}y}{\int_x^{x+1}S(y)\mathrm{d}y}$$

$${}_nm_x=\frac{\int_0^n {}_tp_x\mu_{x+t}\mathrm{d}t}{\int_0^n {}_tp_x\mathrm{d}t}$$

$$= \frac{\int_x^{x+n} S(y)\mu(y)\mathrm{d}y}{\int_x^{x+n} S(y)\mathrm{d}y}$$

6. 随机变量的变换

定理：设连续型随机变量 X 的概率密度函数为 $f(x)$，当 $a < x < b$ 时，$f_X(x) > 0$，$y = g(x)$ 是严格单调增的可微函数，则随机变量 X 的函数 $Y = g(X)$ 的概率密度为：

$$f_Y(y) = \begin{cases} f_X(h(y)) \cdot |h'(y)| & , \quad c < y < d \\ 0 & , \quad \text{其他} \end{cases}$$

其中 $x = h(y)$ 是 $y = g(x)$ 的反函数，$c = \min(g(a), g(b))$，$d = \max(g(a), g(b))$。

如果 $y = g(x)$ 在 (a, b) 内单调递减，仍有反函数 $x = h(y)$ 存在，且 $c = g(b) < y < g(c) = d$ 时，$h(y) < 0$，则 Y 的分布函数为：

$$\begin{aligned} F_Y(y) &= P(Y \leqslant y) \\ &= P(g(X) \leqslant y) \\ &= P(X \geqslant h(y)) \\ &= 1 - P(X < h(y)) \\ &= S(h(y)) \end{aligned}$$

$$\begin{aligned} \lambda_Y(y) &= \frac{f_Y(y)}{S_Y(y)} \\ &= \frac{F'_Y(y)}{1 - S_X(h(y))} \\ &= \frac{S'_X(h(y))}{S_X(h(y))} \cdot \frac{S_X(h(y))}{1 - S_X(h(y))} \cdot h'(y) \\ &= -\lambda_X(h(y)) \cdot \frac{S_X(h(y))}{1 - S_X(h(y))} \cdot \frac{\mathrm{d}x}{\mathrm{d}y} \end{aligned}$$

关于 $E(g(X))$ 的数学期望的讨论：

设 $E(X) = m$，并且 $y = g(x)$ 二次可微，则有：

$$E(g(X)) \approx g(m) + \frac{1}{2}g''(m)\mathrm{Var}(X)$$

$$\mathrm{Var}(g(X)) \approx [g'(m)]^2 \cdot \mathrm{Var}(X)$$

对于二元或多元函数，有类似的近似公式。例如：

$X = g(X_1, X_2), E(X_1) = m_1, E(X_2) = m_2, \mathrm{Var}(X_1) = \delta_1^2$, $\mathrm{Var}(X_2) = \delta_2^2, \mathrm{Cov}(X_1, X_2) = C$，用统计微分法有：

$$E(X) = E(g(X_1, X_2)) \approx g(m_1, m_2)$$

$$\begin{aligned}
\mathrm{Var}(X) &= \mathrm{Var}(g(X_1, X_2)) \\
&\approx (g'_1(m_1, m_2))^2\mathrm{Var}(X_1) \\
&\quad + (g'_2(m_1, m_2))^2\mathrm{Var}(X_2) \\
&\quad + 2g'_1(m_1, m_2) \cdot g'_2(m_1, m_2)\mathrm{Cov}(X_1, X_2) \\
&= (g'_1(m_1, m_2))^2\delta_1^2 + (g'_2(m_1, m_2))^2\delta_2^2 \\
&\quad + 2g'_1(m_1, m_2)g'_2(m_1, m_2) \cdot C
\end{aligned}$$

二、重点、难点解析及例题

本章的重点是生存模型、危险率函数、中心死亡率的概念以及上截尾分布和它的数字特征，解析型的生存模型也是本章的重点，另外，随机变量函数的分布及危险率函数也值得读者关注。本章的难点是危险率函数的定义，下面逐一说明。

所谓危险率函数 μ_x 就是 x 年龄时点上的瞬时死亡水平，也就是在生存到时间 x 的条件下，在时间 x 处的瞬时死亡密度，定义为：

$$\lambda(x) = \mu_x = \lim_{h \to 0}\frac{_hq_x}{h}$$

我们对此式进行整理，整理的过程可使读者更深入地理解 $\lambda(x)$ 或 μ_x。

$$\lambda(x) = \mu_x = \lim_{h \to 0}\frac{_hq_x}{h}$$

$$=\lim_{h\to 0}\frac{P(T(x)<h)}{h}$$

$$=\lim_{h\to 0}\frac{p(x<X<h+x|X>x)}{h}$$

$$=\lim_{h\to 0}\frac{S(x)-S(h+x)}{h\cdot S(x)}$$

$$\underline{\underline{\text{用洛比塔法则}}}\lim_{h\to 0}\frac{-S'(x+h)}{S(x)}$$

$$=\frac{-S'(x)}{S(x)}$$

$$=\frac{f(x)}{S(x)}$$

下面给出几道有关的例题：

例 1 设某随机变量 X 的生存函数 $S(x)=ax^3+b, 0\leqslant x\leqslant k$，若 X 的数学期望为 90，则 X 的方差为(　　)。

A. 90　　B. 180　　C. 360　　D. 450　　E. 540

解：因为，$S(0)=1=ax0^3+b$

所以 $b=1$

因为 $S(k)=0=ak^3+b=a\cdot k^3+1$

所以，$a=-\frac{1}{k^3}$

所以，$f(x)=-S'(x)=\left(\frac{x^3}{k^3}\right)'=\frac{3x^2}{k^3}$

$$\int_0^k x\cdot\frac{3x^2}{k^3}\mathrm{d}x=90$$

$$\Rightarrow\frac{3}{4}k=90$$

$$\Rightarrow k=120$$

所以，$\mathrm{Var}(X)=E(X^2)-90^2$

$$=\frac{3}{5}\cdot 120^2-90^2=540$$

选 E。

例 2 设 X_1 与 X_2 是两个相互独立的随机变量，如果 $Z=\max(X_1,X_2)$，$Y=\min(X_1,X_2)$，则以下选项错误的是( )。

A. Y 的生存函数是 X_1 与 X_2 生存函数的乘积

B. 若 X_1 与 X_2 都服从指数分布，则 Y 也服从指数分布

C. 若 X_1 与 X_2 都服从指数分布，则 Z 不服从指数分布

D. Z 的分布函数为 X_1 与 X_2 分布函数的乘积

E. Z 的密度函数为 X_1 与 X_2 密度函数的乘积

解：
$$\begin{aligned}S_Y(y)&=P(Y>y)=P(\min(X_1,X_2)>y)\\&=P(X_1>y,X_2>y)=P(X_1>y)\cdot p(X_2>y)\\&=S_{X_1}(y)\cdot S_{X_2}(y)\end{aligned}$$

故 A 选项正确。

$$\begin{aligned}F_z(y)&=P(Z\leqslant y)=P(\max(X_1,X_2)\leqslant y)\\&=P(X_1\leqslant y,X_2\leqslant y)=P(X_1\leqslant y)P(X_2\leqslant y)\\&=F_{X_1}(y)\cdot F_{X_2}(y)\end{aligned}$$

故 D 选项正确。

设 $X_1\sim e(\lambda_1)$，$X_2\sim e(\lambda_2)$ 则有：

$$\begin{aligned}S_Y(y)&=S_{X_1}(y)S_{X_2}(y)\\&=\mathrm{e}^{-\lambda_1 y}\cdot\mathrm{e}^{-\lambda_2 y}=\mathrm{e}^{-(\lambda_1+\lambda_2)y}\end{aligned}$$

则 $Y\sim \mathrm{e}(\lambda_1+\lambda_2)$

故 B 选项正确。

$$\begin{aligned}F_Z(y)&=F_{X_1}(y)F_{X_2}(y)=(1-\mathrm{e}^{-\lambda_1 y})(1-\mathrm{e}^{-\lambda_2 y})\\&\neq 1-\mathrm{e}^{-\lambda y}\end{aligned}$$

故 C 选项正确。

选 E。

例 3 已知某随机变量 X 的生存函数为 $S(x)=ax^3+b$，且 $0\leqslant x\leqslant k$，并有 $E(X)=90$，则 ${}_3m_4=$( )。

A. 0.000 054 B. 0.000 027 C. 0.000 013