具有可靠性增长的几何、负二项概型参数的Bayes估计

吴　绍　敏

（数　学　系）

摘　　要

本文研究具有可靠性增长的几何、负二项概型参数的Bayes估计，主要结果是定理1、2，先验分布是均匀的情况作为定理的系给出。

一、引　　言

可靠性测试中，有几何与负二项概型；例如，（一）对一门火炮"膛炸"测试，假定每发射一次成功（不"膛炸"）的概率为p，失败（"膛炸"）的概率为$(1-p)$，若发射X次成功，第$X+1$次失败，此时测试无法继续下去，X是随机应变量，服从几何分布$p(X=x)=p^x(1-p)$，$(x=0, 1, 2\cdots)$，（二），如果取r门同型号、同批生产的火炮进行"膛炸"测试，对每门炮来说，其成功或失败的概率都一样是p或$(1-p)$，假定第i门炮测试成功次数为X_i，则X_i服从几何分布$P(X_i=x_i)=p^{x_i}(1-p)$，$(x_i=0, 1, 2\cdots)$，$i=\overline{1,r}$，且X_1，X_2，…，X_r是相互独立的，则r门炮测试成功的总次数$X=\sum_{i=1}^{r}X_i$服从负二项分布$P(X=k)=\binom{r+k-1}{r-1}p^k(1-p)^r$，$(k=0, 1, \cdots)$。如何对以上两个问题中的$p$进行估计是值得研究的问题。

下面考虑一般的情况。假定系统的可靠性按负二项或几何概型分m级进行测试，第i级测试时，成功概率为$p_i(i=\overline{1,m})$，测试成功次数为X_i，第$i-1$级测试后，经过可靠性改善，第i级测试时的成功概率比p_{i-1}有所增长，本文的目的是应用Bayes方法估计最后一级测试时的成功概率p_m，严格地表达之为下面的定理。

二、p_m的Bayes估计

定理1　假定

(1)$0<p_1<p_2<\cdots p_m<1$

(2)参数 p_i 的先验分布为 Beta 分布

$$B(p_i|a_i,\ b_i)=\frac{1}{B(a_i,\ b_i)}p_i^{a_i-1}(1-p_i)^{b_i-1},\quad a_i,\ b_i \text{ 是正整数},\ i=\overline{1,m}.$$

(3)各级的测试次数 X_i 服从负二项分布

$$f(x_i|p_i,\ r_i)=\binom{r_i+x_i-1}{r_i-1}p_i^{x_i}\left(1-p_i\right)^{r_i},\quad r_i \text{ 为正整数},\ i=\overline{1,m},$$

且 $X_1,\ X_2,\ \cdots,\ X_m$ 相互独立，则在二次损失为 $L(\hat{p}_m,\ p_m)=(\hat{p}_m-p_m)^2$ 的条件下，p_m 的 Bayes 估计为

$$\hat{p}_m=\frac{\sum\limits_{h_1=f_1}^{g_1}\sum\limits_{h_2=f_2}^{g_2}\cdots\sum\limits_{h_{m-1}=f_{m-1}}^{g_{m-1}}W(h_1,\ \cdots,\ h_{m-1})\left(\frac{x_m+a_m+h_{m-1}}{x_m+r_m+a_m+b_m+h_{m-1}}\right)}{W_m},\tag{1}$$

其中

$$\begin{cases} u_i=x_i+a_i, \quad v_i=r_i+b_i, \\ g_i=u_i+v_i+g_{i-1}-1, \quad g_0=0, \\ C_{h_i}=\binom{g_i}{h_i}B\left(u_i+h_{i-1},\ g_{i-1}+v_i-h_{i-1}\right), \quad (i=\overline{1,m}) \\ f_{i-1}=u_{i-1}+h_{i-2}, \quad h_0=0,\ f_0=0, \\ W(h_1,\ h_2,\ \cdots,\ h_{m-1})=\left(\prod\limits_{i=1}^{m-1}C_{h_i}\right)B\left(u_m+h_{m-1},\ g_{m-1}+v_m-h_{m-1}\right), \\ W_m=\sum\limits_{h_1=f_1}^{g_1}\sum\limits_{h_2=f_2}^{g_2}\cdots\sum\limits_{h_{m-1}=f_{m-1}}^{g_{m-1}}W(h_1,\ h_2,\ \cdots,\ h_{m-1}). \end{cases}\tag{2}$$

证明 记 $P=(p_1,\ p_2,\ \cdots,\ p_m)$， $X=(X_1,\ X_2,\ \cdots,\ X_m)$， $r=(r_1,\ r_2,\ \cdots,\ r_m)$，

$$D_m=\left\{(p_1,\ p_2,\ \cdots,\ p_m):0<p_1<p_2<\cdots<p_m<1\right\},$$

在 p，r 固定下 X 的条件概率密度为

$$\begin{aligned} f(x|p,\ r)&=\prod_{i=1}^{m}f(x_i|p_i,\ r_i)=\prod_{i=1}^{m}\binom{r_i+x_i-1}{r_i-1}p_i^{x_i}(1-p_i)^{r_i} \\ &=\left\{\prod_{i=1}^{m}\binom{r_i+x_i-1}{r_i-1}\right\}\prod_{i=1}^{m}p_i^{x_i}(1-p_i)^{r_i}. \end{aligned}\tag{3}$$

$P=(p_1,\ p_2,\ \cdots,\ p_m)$ 的联合分布密度为

$$\pi(P)=\begin{cases} \dfrac{\prod\limits_{i=1}^{m}p_i^{a_i-1}(1-p_i)^{b_i-1}}{\int\cdots\int\limits_{D_m}\prod\limits_{i=1}^{m}p_i^{a_i-1}(1-p_i)^{b_i-1}\,dp_1\cdots dp_m}, & \text{当 } P\in D_m \\ 0, & \text{当 } P\overline{\in}D_m \end{cases}\tag{4}$$

$$f(p|x,r)=\frac{f(x|p,\ r)\pi(p)}{\int\cdots\int\limits_{D_m}f(x|p,\ r)\pi(p)\,dp_1,\ dp_2\cdots dp_m}=\frac{\prod\limits_{i=1}^{m}p_i^{x_i+a_i-1}(1-p_i)^{r_i+b_i-1}}{W_m}$$

$$\overset{(2)}{=} \frac{\prod_{i=1}^{m} p_i^{u_i-1}(1-p_i)^{v_i-1}}{W_m}, \tag{5}$$

其中

$$W_m = \int\cdots\int_{D_m} \prod_{i=1}^{m} p_i^{u_i-1}(1-p_i)^{v_i-1} dp_1\cdots dp_m .$$

于是 p_m 的条件密度为

$$f(p_m|x,\ r) = \frac{p_m^{u_m-1}(1-p_m)^{v_m-1}\int_0^{p_m} p_{m-1}^{u_{m-1}-1}(1-p_{m-1})^{v_{m-1}-1}dp_{m-1}\cdots\int_0^{p_1} p_1^{u_1-1}(1-p_1)^{v_1-1}dp_1}{W_m} \tag{6}$$

再由 Beta 分布的性质知，当 u、v 都是正整数时有恒等式

$$\int_0^y x^{u-1}(1-x)^{v-1}dx = B(u,\ v)\sum_{j=u}^{u+v-1}\binom{u+v-1}{j} y^j(1-y)^{u+v-1-j} \tag{7}$$

反复应用(7)式，可计算(6)式的分子

$$\begin{aligned}
&p_m^{u_m-1}(1-p_m)^{v_m-1}\int_0^{p_m} p_{m-1}^{v_{m-1}-1}(1-p_{m-1})^{v_{m-1}-1}dp_{m-1}\cdots\int_0^{p_3} p_2^{u_2-1}(1-p_2)^{v_2-1}\\
&\quad\times\left\{\sum_{h_1=u_1}^{u_1+v_1-1}\binom{u_1+v_1-1}{h_1}B(u_1,\ v_1)p_2^{h_1}(1-p_2)^{u_1+v_1-1-h_1}\right\}dp_2\\
&=p_m^{u_m-1}(1-p_m)^{v_m-1}\int_0^{p_m} p_{m-1}^{u_{m-1}-1}(1-p_{m-1})^{v_{m-1}-1}dp_{m-1}\cdots\\
&\quad\times\left\{\sum_{h_1=u_1}^{g_1} C_{h_1}\int_0^{p_3} p_2^{u_2+h_1-1}(1-p_2)^{g_1+v_2-1-h_1}dp_2\right\}\\
&=p_m^{u_m-1}(1-p_m)^{v_m-1}\int_0^{p_m} p_{m-1}^{u_{m-1}-1}(1-p_{m-1})^{v_{m-1}-1}dp_{m-1}\cdots\\
&\quad\times\int_0^{p_4} p_3^{u_3-1}(1-p_3)^{v_3-1}\left\{\sum_{h_1=u_1}^{g_1} C_{h_1}\sum_{h_2=u_2+h_1}^{g_2} C_{h_2}p_3^{h_2}(1-p_3)^{g_2-h_2}\right\}dp_3\\
&=p_m^{u_m-1}(1-p_m)^{v_m-1}\int_0^{p_m} p_{m-1}^{u_{m-1}-1}(1-p_{m-1})^{v_{m-1}-1}dp_{m-1}\cdots\\
&\quad\times\int_0^{p_5} p_4^{u_4-1}(1-p_4)^{v_4-1}\left\{\sum_{h_1=f_1}^{g_1}\sum_{h_2=f_2}^{g_2} C_{h_1}C_{h_2}\int_0^{p_4} p_3^{u_3+h_2-1}(1-p_3)^{g_3+v_3-1-h_2}dp_3\cdots\right.\\
&=\sum_{h_1=f_1}^{g_1}\sum_{h_2=f_2}^{g_2}\cdots\sum_{h_{m-1}=f_{m-1}}^{g_{m-1}} C_{h_1}\cdots C_{h_{m-1}}p_m^{u_m+h_{m-1}-1}(1-p_m)^{g_{m-1}+v_m-1-h_{m-1}}\\
&=\sum_{h_1=f_1}^{g_1}\sum_{h_2=f_2}^{g_2}\cdots\sum_{h_{m-1}=f_{m-1}}^{g_{m-1}} W(h_1,h_2,\cdots,h_{m-1})B(p_m|u_m+h_{m-1},g_{m-1}+v_m-h_{m-1}).
\end{aligned} \tag{8}$$

故有

$$f(p_m|x,\ r)=\frac{\sum\limits_{h_1=f_1}^{g_1}\sum\limits_{h_2=f_2}^{g_2}\cdots\sum\limits_{h_{m-1}=f_{m-1}}^{g_{m-1}}W(h_1,h_2,\cdots,h_{m-1})B(p_m|u_m+h_{m-1},g_{m-1}+v_m-h_{m-1})}{W_m}. \tag{9}$$

(8)式两边对 p_m 从 0 到 1 积分即得(2)式中的 W_m。在二次损失为 $L(\hat{P}_m,\ P_m)=(\hat{P}_m-P_m)^2$，$P_m$ 的 Bayes 估计是

$$\hat{P}_m=E(P_m|X)=\int_0^1 p_m f(p_m|x,\ r)dP_m. \tag{10}$$

因

$$\int_0^1 xB(x|u,\ v)\ dx=\frac{u}{u+v}, \tag{11}$$

故由（9），（10），（11）算得（1）式的 $\hat{P}_m$，证毕。

特别，当 $a_i=b_i=1$，（$i=\overline{1,m}$）时，得

系 1：当 $P_i(i=\overline{1,m})$的先验分布为$U(0,\ 1)$，则

$$\pi(P)=\begin{cases}m!\ , & 当P\in D_m.\\ 0\ , & 当P\,\overline{\in}\,D_m.\end{cases} \tag{12}$$

$$\hat{P}_m=\frac{\sum\limits_{h_1=f_1}^{g_1}\sum\limits_{h_2=f_2}^{g_2}\cdots\sum\limits_{h_{m-1}=f_{m-1}}^{g_{m-1}}W(h_1,\ \cdots,\ h_{m-1})\left(\dfrac{x_m+h_{m-1}+1}{x_m+g_{m-1}+r_m+2}\right)}{W_m}, \tag{13}$$

其中

$$\begin{cases}u_i=x_i+1,\ \ v_i=r_i+1, & i=\overline{1,m},\\ g_i=u_i+v_i+g_{i-1}-1=x_i+r_i+g_{i-1}+1, & g_0=0,\\ C_{h_i}=\dbinom{g_i}{h_i}B(u_i+h_{i-1},\ \ g_{i-1}+v_i-h_{i-1}), & h_0=0,\\ f_{i-1}=x_{i-1}+1+h_{i-2}, & f_0=0.\end{cases} \tag{14}$$

定理 2　如果测试是按几何概型，即满足

（1）$0<P_1<P_2<\cdots<P_m<1$，

（2）参数 P_i 的先验分布为 Beta 分布

$$B(p_i|a_i,\ b_i)=\frac{1}{B(a_i,\ b_i)}p_i^{a_i-1}\ (1-p_i)^{b_i-1}\qquad a_i,\ b_i 是正整数,$$

（3）各级的测试次数 X_i，服从几何分布

$$f(x_i|p_i)=p_i^{x_i}(1-p_i).\qquad\qquad i=\overline{1,m}\ .$$

且 X_1，X_2，…，X_m 相互独立，则在二次损失下 P_m 的 Bayes 估计为

$$\hat{P}_m=\frac{\sum\limits_{h_1=f_1}^{g_1}\sum\limits_{h_2=f_2}^{g_2}\cdots\sum\limits_{h_{m-1}=f_{m-1}}^{g_{m-1}}W(h_1,\ \cdots,\ h_{m-1})\left(\dfrac{x_m+a_m+h_{m-1}}{x_m+a_m+b_m+g_{m-1}+1}\right)}{W_m} \tag{15}$$

其中

$$\begin{cases} u_i = x_i + a_i, \quad v_i = b_i + 1 \quad (i = \overline{1, m}) \\ g_i = x_i + a_i + b_i + g_{i-1}, \quad g_0 = 0 \\ C_{h_i} = \binom{g_i}{h_i} B(x_i + a_i + h_{i-1}, \ g_{i-1} + b_i - h_{i-1} + 1), \quad h_0 = 0 \\ f_{i-1} = x_{i-1} + a_{i-1} + h_{i-2}, \quad f_0 = 0 \\ W(h_1, \ h_2, \ \cdots h_{m-1}) = \left(\prod_{i=1}^{m-1} C_{h_i}\right) B(x_m + a_m + h_{m-1}, \ g_{m-1} + b_m + 1 - h_{m-1}), \\ W_m = \sum_{h_1=f_1}^{g_1} \sum_{h_2=f_2}^{g_2} \cdots \sum_{h_{m-1}=f_{m-1}}^{g_{m-1}} W(h_1, \ h_2, \ \cdots, \ h_{m-1}), \end{cases}$$

证　因几何分布是负二项分布的特殊情况，故只要在定理 1 中令 $r_i = 1$，$(i = \overline{1, m})$，即得定理 2。

系 2：当 $P_i (i = \overline{1, m})$ 的先验分布为 $U(0, 1)$ 则

$$\pi(P) = \begin{cases} m!, & \text{当} P \in D_m, \\ 0, & \text{当} P \bar{\in} D_m, \end{cases}$$

$$\hat{P}_m = \frac{\sum_{h_1=f_1}^{g_1} \sum_{h_2=f_2}^{g_2} \cdots \sum_{h_{m-1}=f_{m-1}}^{g_{m-1}} W(h_1, \ h_2, \ \cdots, \ h_{m-1}) \left(\frac{x_m + h_{m-1} + 1}{x_m + g_{m-1} + 3}\right)}{W_m}, \tag{16}$$

其中

$$\begin{cases} u_i = x_i + 1, \quad v_i = 2, \quad (i = \overline{1, m}), \\ g_i = x_i + g_{i-1} + 2, \quad g_0 = 0 \\ C_{h_i} = \binom{g_i}{h_i} B(x_i + h_{i-1} + 1, \ g_{i-1} - h_{i-1} + 2), \\ f_{i-1} = x_{i-1} + h_{i-2} + 1, \quad h_0 = 0, \quad f_0 = 0, \\ W(h_1, \ h_2, \ \cdots, \ h_{m-1}) = \left(\prod_{i=1}^{m-1} C_{h_i}\right) B(x_m + h_{m-1} + 1, \ g_{m-1} + 2 - h_{m-1}), \\ W_m = \sum_{h_1=f_1}^{g_1} \sum_{h_2=f_2}^{g_2} \cdots \sum_{h_{m-1}=f_{m-1}}^{g_{m-1}} W(h_1, \ h_2, \ \cdots, \ h_{m-1}), \end{cases}$$

证：只要在定理 2 中令 $a_i = b_i = 1$，$(i = \overline{1, m})$，即得。

顺便提一下 P_m 的置信下限 $p_{m, L}$ 可由下式确定

$$P\left\{P_m \geqslant p_{m, L}\right\} = \int_{p_{m, L}}^{1} f(p_m \mid x, \ r) \, dp_m = 1 - \alpha. \tag{17}$$

参考文献

〔1〕陈世基，《具有可靠性增长的二、三项分布概型参数 Bayes 估计的注记》，福建师大，自然科学版，1983，第 1 期。

〔2〕〔美国〕，N. R. 曼，R. E. 沙费尔，N. D. 率普瓦拉，《可靠性与寿命数据的统计分析方法》，1980，12。