第 35 卷第 1 期　　湖北工业大学学报　　2020 年 02 月
Vol.35 No.1　　Journal of Hubei University of Technology　　Feb.2020

[文章编号] 1003－4684(2020)01-0001-04

基于假设检验法的插秧机寿命建模

文昌俊[1,2]，王　冕[1,2]，陈　立[1,2]

(1 湖北工业大学机械工程学院，湖北 武汉 430068；2 湖北省现代制造质量工程重点实验室，湖北 武汉 430068)

[摘　要] 可靠性差、故障多和寿命周期难以准确估算是当前国产插秧机普遍存在的问题，而当前研究插秧机寿命及寿命建模的相关成果较少。以某型插秧机为研究对象，收集并统计该型插秧机的故障时间数据，使用假设检验法作为建立插秧机寿命模型的理论依据。假设该型插秧机寿命分布为二参数威布尔模型，根据故障数据，运用χ^2检验法进行模型检验，而后分别使用最小二乘法和极大似然法对模型参数进行估计，并对估计值进行优选。运用 K-S 检验法进行拟合优度检验，以检验模型的显著性。

[关键词] 插秧机寿命模型；威布尔分布；参数估计；拟合优度检验

[中图分类号] S223.91　　[文献标识码] A

与进口插秧机相比，国产插秧机性能基本满足需求，但故障次数较多，寿命周期较短，而适用于农业机械产品可靠性与寿命预测、评估的理论和方法还不够成熟。虽然电子类产品的可靠性与寿命研究比较成熟，该行业建立了完善的可靠性标准，可靠性与寿命评估方法规范系统，但农业机械毕竟与电子产品区别较大，不能完全照搬其方法[1]。针对现状，陈建能等对水稻插秧机可靠性评价指标进行改进，提出用维护系数作为衡量其可靠性的一个指标[2]。郭清臣等用随机过程理论对拖拉机可靠性进行分析[3]。本文依据故障样本数据，建立了插秧机寿命分布模型，检验通过，证明了模型的适用性。

1　插秧机寿命模型建立

1.1　机械可靠性寿命模型识别方法

在机械可靠性与寿命评估时，确定寿命分布类型具有重要意义。常用识别寿命模型的方法有图形法、假设检验法、回归分析法和基于失效机理分析法。相比其他三种方法来说，假设检验法简单易行，识别模型效果优良，所以本文采用假设检验法对插秧机寿命模型做出选择。

1.2　插秧机寿命模型确定

1.2.1　威布尔分布模型的选择　机械产品常用的寿命分布有正态分布、威布尔分布、指数分布和对数分布，而威布尔分布是其中最常用的一种。依前所述，假设插秧机的寿命分布为威布尔分布，而后运用χ^2检验法对分布进行检验，威布尔模型参数确定后使用 K-S 检验法对分布进行拟合优度检验，以进一步确保寿命分布模型的准确性。

威布尔分布有三参数和二参数两种形式。当其位置参数 $\gamma\neq0$ 时，为三参数威布尔分布；当位置参数 $\gamma=0$ 时，则为二参数威布尔分布。

插秧机的寿命分布用威布尔模型进行拟合，则模型的随机变量为插秧机的寿命 t。插秧机从零时刻开始作业，故障产生是随机的，所以从开始工作就可能出现故障，故威布尔模型的位置参数为零，即 $\gamma=0$。由此可得：插秧机的寿命模型为二参数威布尔分布，相应的故障概率密度函数和故障概率分布函数分别为：

$$f(t)=\frac{m}{\eta}\left(\frac{t}{\eta}\right)^{m-1}\exp\left(-\left(\frac{t}{\eta}\right)^{m}\right)\tag{1}$$

$$F(t)=1-\exp\left(-\left(\frac{t}{\eta}\right)^{m}\right)\tag{2}$$

式中：m 为形状参数；η 为尺度参数；t 为插秧机故障时间。

1.2.2　威布尔分布模型的检验　运用 χ^2 检验法对插秧机的寿命模型是否为威布尔分布进行检验。假设某产品的寿命分布为 $F(t)$，检验假设 H_0：

$$F(t)=F_0(t;\eta,m)=1-\exp[-(t/\eta)^m]$$

式中：m、η 是未知参数。

任意取 n 个产品进行寿命实验，到 r 个发生故障时实验停止，故障时间依次 $t_{(1)}\leqslant t_{(2)}\leqslant\cdots\leqslant t_{(r)}$。设

[收稿日期] 2019－07－16

[第一作者] 文昌俊(1970－)，男，湖北仙桃人，湖北工业大学教授，研究方向为可靠性与质量控制

[通信作者] 王　冕(1994－)，男，湖北十堰人，湖北工业大学硕士研究生，研究方向为可靠性与质量控制

$X_i = lnt_i$，若令

$$V_i = (r-i)(X_{(r-i+1)} - X_{(r-i)}), (i=1,2,\cdots,r-1) \tag{3}$$

于是 $V_i/\sigma(i=1,2,\cdots,r-1)$是近似相互独立，且同时服从标准指数分布的随机变量。根据指数分布的性质，可以知道 $2V_i/\sigma(i=1,2,\cdots,r-1)$都是近似独立，且服从具有自由度为 2 的$\chi^2$分布。由巴特例特统计量可得：

$$B^2 = 2(r-1)\lg(\sum_{i=1}^{r-1} V_i/(r-1)) - 2\sum_{i=1}^{r-1}\lg V_i \tag{4}$$

$$c = 1 + \frac{r}{6(r-1)} \tag{5}$$

且 B^2/c 是自由度为$r-2$ 的χ^2变量。

如果给定显著性水平α，由χ^2分布表可查得$\chi^2_{\alpha/2}(r-2)$和$\chi^2_{1-\alpha/2}(r-2)$，当 $B^2/c<\chi^2_{\alpha/2}(r-2)$或$B^2/c>\chi^2_{1-\alpha/2}(r-2)$时，拒绝假设 H_0；否则，接受H_0。

根据某企业一个作业季的某型号插秧机维修记录和现场故障跟踪试验中，采集到 80 条插秧机故障时间数据：16.0、16.4、29.3、31.6、…、325.1、325.8 h。故障时间依次为 t_1，t_2，…，t_{80}，由式(3)可计算得 V_i(表 1)。

表 1 V_i结果

序号 i	V_i	序号 i	V_i
1	0.16992	…	…
2	13.70932	79	0.02469

由式(4)、(5)可得：$B^2=78.18361$，$C=1.16878$，$B^2/C=66.8936$。取显著性水平$\alpha=0.05$，由χ^2分布表，查得$\chi^2_{0.025}(78)=55.466$、$\chi^2_{0.975}(78)=104.316$。由 $55.466<B^2/c=66.8936<104.316$，即$\chi^2_{\alpha/2}(r-2)<B^2/c<\chi^2_{1-\alpha/2}(r-2)$，因此接受假设 H_0，即该插秧机的寿命分布模型为二参数威布尔分布。

2 插秧机寿命模型参数估计

插秧机的寿命模型为二参数威布尔分布，依据文献[1]所述的参数估计法，采用最小二乘估计和极大似然估计两种方法对模型的参数进行估计，以确保参数的精确度。

2.1 最小二乘估计

假设某型插秧机在观测周期内共发生 n 次故障，其故障发生时间依次为 $t_1<t_2<\cdots<t_n$，若该型插秧机寿命 T 服从二参数威布尔分布，即 $T\sim W(m,\eta)$。

将数据$\{(t_i,F(t_i))\}(i=1,2,\cdots,n)$代入式(2)并对其进行线性化，令 $Y_i=lnln[1/(1-F(t_i))]$，$X_i=lnt_i$，$A=-mln(\eta)$，$B=m$，则参数 m、η、A和 B 的估计值表达式为：

$$\hat{m} = \hat{B} = \frac{\sum_{i=1}^{n}(X_iY_i) - \frac{1}{N}(\sum_{i=1}^{n}X_i)(\sum_{i=1}^{n}Y_i)}{\sum_{i=1}^{n}X_i^2 - \frac{1}{N}(\sum_{i=1}^{n}X_i)^2} \tag{6}$$

$$\hat{A} = \frac{1}{N}\sum_{i=1}^{n}Y_i - \frac{B}{N}\sum_{i=1}^{n}X_i \tag{7}$$

$$\hat{\eta} = \exp\left(-\frac{\hat{A}}{\hat{B}}\right) \tag{8}$$

设 $MR(t_i)$为 t_i时刻的中位秩，用中位秩 $MR(t_i)$作为 $F(t_i)$的估计值，公式如下：

$$MR(t_i)=F(t_i)=[n(t_i)-0.3]/(n+0.4) \tag{9}$$

依据插秧机故障发生时间数据 $t_i(i=1,2,\cdots,80)$，分别计算出 $F(t_i)$，X_i和 Y_i(表 2)。

表 2 插秧机故障数据表

序号 i	插秧机寿命 t_i	中位秩 $F(t_i)$	X_i	Y_i
1	16.0	0.0087	2.7726	−4.7393
2	16.4	0.0211	2.7973	−3.8457
⋮	⋮	⋮	⋮	⋮
79	325.1	0.9789	5.7841	1.3497
80	325.8	0.9913	5.7863	1.5568

将表 1 数据代入式(6)、(7)和(8)，可计算出：

$$\hat{m}=1.7744, \hat{\eta}=184.3894, \hat{A}=-9.2572, \hat{B}=1.7744$$

2.2 极大似然估计

把 $t_1,t_2,\cdots,t_n$代入式(3)，并求自然对数，得到似然函数如下：

$$L(m,\eta) = -m\eta\ln\eta + n\ln m + (m-1)\sum_{i=1}^{n}\ln t_i - \frac{1}{\eta^m}\sum_{i=1}^{n}t_i)^m \tag{10}$$

对式(10)取关于参数 m，η 的偏导数，得到似然方程组：

$$\begin{cases} \frac{\partial L}{\partial m} = -n\ln\eta + \frac{n}{m} + \sum_{i=1}^{n}\ln t_i + \\ \quad \frac{\ln\eta}{\eta^m}\sum_{i=1}^{n}t_i)^m - \frac{1}{\eta^m}\sum_{i=1}^{n}t_i)^m\ln t_i = 0 \\ \frac{\partial L}{\partial \eta} = -\frac{mn}{\eta} + m\frac{1}{\eta^{m+1}}\sum_{i=1}^{n}t_i)^m = 0 \end{cases} \tag{11}$$

基于 MATLAB 编程，运用遗传算法对式(11)似然方程组求解，得：

$$\hat{m}=2.2245, \hat{\eta}=179.1257$$

由上述结果可知，最小二乘估计和极大似然估计两种方法分别得到的威布尔模型参数估计值间存在一定差距，需对两种方法解出的参数估计值进行比较与优选。

2.3 参数估计值优选

1)估计量优选的评定标准

根据文献[10]中估计量优良性的评定标准，选

用无偏性和有效性作为参数估计值的优选标准。

2)最小二乘估计得出威布尔分布的数学期望

$$E_z(t) = 164.094$$

使用 MINITAB 软件计算得出该分布的标准差σ_z为 95.5656。

3)极大似然估计得出威布尔分布的数学期望

$$E_j(t) = 158.646$$

使用 MINITAB 软件计算得出该分布的标准差σ_j为 75.3695。

4)样本均值和标准差分别与两种估计法得到的数学期望与标准差对比分析

由插秧机故障样本数据，得出样本均值 $\bar{t}=159.386$，样本标准差 $s=75.4378$。样本均值分别与最小二乘法和极大似然法估计得出的模型数学期望比较见表 3。

表 3　样本均值分别与最小二乘和极大似然估计的模型数学期望比较

	$\frac{\lvert \bar{t}-E_z(t)\rvert}{E_z(t)}$	$\frac{\lvert \bar{t}-E_j(t)\rvert}{E_j(t)}$
比较值	2.869%	0.466%

由表 3 得：故障样本均值更接近极大似然估计得出的威布尔模型的数学期望。样本标准差分别与最小二乘法和极大似然法估计得出的模型标准差比较见表 4。

表 4　样本标准差分别与最小二乘和极大似然估计的模型标准差比较

	$\lvert s-\sigma_z\rvert$	$\lvert s-\sigma_j\rvert$
差值	20.1278	0.0683

极大似然估计得出的模型标准差小于最小二乘估计得出的模型标准差，即$\sigma_j=75.3695<\sigma_z=95.5656$。同时由表 4 得：$\lvert s-\sigma_j\rvert<\lvert s-\sigma_z\rvert$，即插秧机故障样本数据标准差与极大似然估计得出的模型标准差相差较小。

综上所述：极大似然法估计得到的威布尔模型参数值优于最小二乘法估计得到的威布尔模型参数值，更适合于描述该型插秧机的寿命分布情况。即：

$$\hat{m} = 2.2245\ ,\ \hat{\eta} = 179.1257$$

3　插秧机寿命模型拟合优度检验

K-S(Kolmogorov-Smirnov/柯尔莫哥洛夫-斯米尔诺夫)检验法可用来检验一组样本数据的分布与某一指定的理论分布之间的符合程度。

设总体分布为 $F(t)$，$F_0(t)$为已知的连续分布函数，假设检验问题 H_0：$F(t)=F_0(t)$，K-S 检验步骤如下。

1)样本量为 n 的样本 $T_1,T_2,\cdots,T_n$，其顺序统计量为：$T_{(1)}\leqslant T_{(2)}\cdots\leqslant T_{(n)}$，可得到经验分布函数为：

$$F_n(t_i) = \begin{cases} 0, & t\leqslant T(1) \\ \dfrac{i}{n}, & T_i\leqslant t<T_{i+1} \\ 1, & t\geqslant T_n \end{cases} \tag{12}$$

柯尔莫哥洛夫提出检验假设 H_0的统计量

$$L_n = \max_{i=1,\cdots,n}\{\lvert F_n(t_i)-F_0(t_i)\rvert\} \tag{13}$$

2)计算各个样本数据 t_i对应的经验分布函数 $F_n(t_i)$和已知的分布函数 $F_0(t_i)$、并联立式(12)、(13)得：

$$L_n = \max_{i=1,\cdots,n}\left\{\left|F_0 t_i-\frac{i-1}{n}\right|,\left|F_0 t_i-\frac{i}{n}\right|\right\} \tag{14}$$

在计算统计量 L_n时，先求出$\delta_i=\max\{\lvert F_0(t_i)-(i-1)/n\rvert,\lvert F_0(t_i)-i/n\rvert\}$，$(i=1,\cdots,n)$，然后在$\delta_1,\cdots,\delta_n$中选择最大的一个就是 L_n，即 $L_n=\max_i\{\delta_i\}$。

3)对于给定的显著性水平α和样本量 n，由柯尔莫哥洛夫检验临界表查出临界值 $d_{n,\alpha}$。当 $L_n\leqslant d_{n,\alpha}$时，接受假设 H_0；反之拒绝假设 H_0。

假设 H_0：插秧机的故障数据服从 $m=2.2245$，$\eta=179.1257$ 的二参数威布尔分布，即

$$F_0(t_i) = 1-\exp\left(-\left(\frac{t_i}{179.1257}\right)^{2.2245}\right) \tag{15}$$

依据式(14)、(15)，使用 Matlab 编程解出 $F_0(t_i)$和δ_i(表 5)。

表 5　插秧机寿命分布的检验

i	t_i/h	$(i-1)/n$	i/n	δ_i
1	16.0	0	0.0125	0.007871842
2	16.4	0.0125	0.0250	0.020111151
⋮	⋮	⋮	⋮	⋮
80	325.8	0.9875	1.0000	0.022739744

从表 5 可得：$L_n=\max_i\{\delta_i\}=0.141788$。

由显著性水平$\alpha=0.05$，样本量 $n=80$，通过查表 $d_{n,\alpha}=0.14960$，$L_n<d_{n,\alpha}$，所以接受假设 H_0，即插秧机的故障数据服从 $m=2.2245$，$\eta=179.1257$ 的二参数威布尔分布。

4　结论

1)结合插秧机现场故障数据建立了二参数威布尔模型。在得出模型参数值前，运用χ^2检验法进行检验，检验通过，证明威布尔模型适合描述该型插秧机的寿命情况。

2)分别使用最小二乘法和极大似然法对模型的参数进行估计，得到两组参数估计值后，使用无偏性和有效性作为估计值的评定标准，对估计值进行

优选，最后得出极大似然法估计的参数值更好。在确定模型参数估计值后，进行分布拟合优度检验，得模型的显著性水平。

[参 考 文 献]

[1] 王艳芳.基于 bootstrap 法与混合威布尔分布的拖拉机可靠性评估模型与应用研究[D],哈尔滨:东北农业大学,2016.

[2] 陈建能,赵匀.水稻插秧机可靠性评价指标的改进[J],农机化研究,2002(4):39-41.

[3] 郭清臣,柳玲文.拖拉机整机可靠性试验数据分析方法[J],洛阳工学院学报，1991,12(2):64-69.

[4] 王亚菲,邬海峰,基于威布尔分布的汽车寿命估计方法[J],江南大学学报,2015,14(2):248-252.

[5] 郝静如.机械可靠性工程[M].北京:国防工业出版社,2006.

[6] 赵宇.可靠性数据分析[M].北京:国防工业出版社,2011.

[7] 罗静,陈一凡,张根宝,等.数控磨床可靠性模型优选与分析[J].现代制造工程，2018(11):119-125.

[8] 陈振中,王航超,王璐璐,等.基于威布尔分布的 APU 二级涡轮叶片可靠性分析[J].沈阳航空航天大学学报,2018,35(5):23-28.

[9] Emad E, Elmahdy. A new approach for Weibull modeling for reliability life data Analysis[J]. Applied Mathematics and Computation, 2015, 250:708-720.

[10] 陈魁.应用概率统计[M].北京:清华大学出版社,2000.

Research on Life Model of Rice Transplanter

WEN Changjun[1,2], WANG Mian[1,2], CHEN Li[1,2]

(1 *School of Mechanical Engineering, Hubei Univ. of Tech., Wuhan* 430068, *China*;

2 *Hubei Key Laboratory of Modern Manufacturing Quality Engineering, Wuhan* 430068, *China*)

Abstract: The poor reliability, multiple faults and difficulty in estimating accurately life cycle are common problems in domestic rice transplanters. However, there are few achievements in the research of transplanter life and life modeling. This paper takes a type of transplanter as the research object, and the fault time data of transplanter is counted. The hypothesis test method is used as the theoretical basis for establishing the life model of the transplanter. It is assumed that the life distribution of the transplanter is a two-parameter Weibull model. According to the fault data, the model is tested by the chi-square test method. Then the least square method and maximum likelihood method are used to estimate the model parameters, and the estimated values are optimized. The k-s test method is used to test the goodness of fit; it is used to test the significance of the model.

Keywords: rice transplanter life model; Weibull distribution; parameter estimation; goodness of fit test

[责任编校: 张 众]