

doi: 10.3969/j.issn.1006-1576.2010.08.011

基于置信区间的信号仿真决策分析模型

胡韬, 胡伟文

(海军工程大学 理学院, 湖北 武汉 430033)

摘要: 为了对信号仿真的决策分析进行研究, 采用置信区间的方法分析信号特征参数的一致性。讨论了仿真输出抽样得到的参数置信区间与给定接收区间的长度匹配问题, 完善了置信区间法; 针对仿真一致性不够高的情况, 建立了置信区间对于接收区间的偏离函数, 提高信号仿真模型的调试效率, 在此基础上提出了基于置信区间的信号仿真决策分析模型。最后, 给出了该模型在武器装备水下航行体辐射噪声仿真模型中的应用实例。结果表明, 该模型能充分挖掘仿真信号抽样所得到的有用信息, 可为信号仿真建模 VV&A 提供辅助决策, 进而提高仿真置信度。

关键词: 仿真; 仿真模型 VV&A; 置信区间; 决策分析模型

中图分类号: TP391.9; C934 **文献标识码:** A

Decision-Making Analysis Model of Signal Simulation Based on Confidence Interval

Hu Tao, Hu Weiwen

(College of Science, Naval University of Engineering, Wuhan 430033, China)

Abstract: To study the decision-making analysis of signal simulation, the method of the confidence interval is adopted to analyze the consistency of signal characteristic parameter. The matching problem of the confidence interval which comes from simulation sampling and the acceptance interval is discussed, which perfects the method of the confidence interval; Then, a deviated function is given about the confidence interval to the acceptance interval, which can increase the debugging efficiency of simulation model. Based on this, a decision-making analysis model of signal simulation based on the confidence interval is put forward. Finally, an example applied in the simulation of underwater target radiated-noise is taken into explanation. It illustrates that the model can scoop out a plentiful of useful information richly, which comes from simulation sampling and can provide auxiliary decision-making for the simulation model VV&A and then increase the credibility of simulation.

Keywords: simulation; simulation model VV&A; confidence interval; decision-making analysis model

0 引言

由于实际信号的采集不仅耗时、耗力而且受到经济、环境等条件的限制, 在水中兵器仿真试验中, 水中目标辐射噪声常部分地采用仿真信号。信号仿真基于模型进行, 其不完全可达性^[1]决定了结果不可能完全一致地反映实际信号的全部特征, 这就存在一个仿真置信度的问题^[2-6]。保证仿真置信度的关键是对仿真模型的验证, 在仿真一致性、目的性之间综合权衡并作出决策。

目前, 仿真置信度的研究多是围绕其评估理论和方法来展开, 文献[3]利用相似度的方法对舰艇鱼雷弹道的仿真置信度进行相似评估, 文献[7]主要根据模型输出数据的置信区间原理建立了模型置信度检验模型, 并对已有装备论证模型进行了置信度检验演算。但当仿真置信度不够高时, 如何提高仿真置信度, 尤其是信号仿真的决策分析的研究还不是

很多。故采用置信区间的方法, 提出了基于置信区间的信号仿真决策分析模型, 为信号仿真建模 VV&A 提供辅助决策。

1 基于置信区间的仿真信号一致性分析

目前, 在仿真模型置信度的研究上已经存在着一定的共识。比如, 在某一特征参数上, 实际系统与仿真模型系统输出的一致性被认为是衡量仿真置信度的重要标准。一般情况下, 对仿真模型进行一致性分析采用的置信度法^[7]是比较模型输出数据所产生的置信区间和给定的接收区间, 进而判断仿真模型的置信度。

对于某仿真信号的一个特征参数 θ 而言, 若对于给定置信水平 $1-\alpha$, 存在 2 个统计量 $T_1 = T_1(x_1, x_2, \dots, x_n)$ 、 $T_2 = T_2(x_1, x_2, \dots, x_n)$ 使得式 (1) 成立, 则称区间 (T_1, T_2) 为参数 θ 的 $1-\alpha$ 置信区间^[8]。

$$p\{T_1(x_1, x_2, \dots, x_n) < \theta < T_2(x_1, x_2, \dots, x_n)\} = 1-\alpha \quad (1)$$

收稿日期: 2010-02-13; 修回日期: 2010-04-23

基金项目: 海军工程大学自然科学基金 (HGDJJ200706)

作者简介: 胡韬 (1983-), 男, 辽宁人, 在读硕士生, 从事武器系统效能评估与信息处理研究。

在单一特征参数的情形下，检验方能够接受的参数的误差范围称之为接收区间^[9]。而参数区间估计的意义：置信度是被估计参数落入置信区间的概率，或者说置信区间以等于置信度的概率包含被估计参数。用置信区间与接收区间的关系来分析仿真信号的一致性。为行文方便，定量分析置信区间与接收区间的关系，引用文献[9]的随机数如下：

$$\xi = \frac{[a_i, b_i] \cap [a, b] \text{ 的长度}^{[9]}}{[a_i, b_i] \cup [a, b] \text{ 的长度}} \quad (2)$$

随机数 $\xi \in [0, 1]$ ，其大小反映区间之间的重叠程度。 ξ 的值越大，通过仿真抽样得到的置信区间与接收区间的重叠程度就越高，仿真信号的一致性就越强；反之，则一致性就越弱。因此，可通过 ξ 来定量分析仿真信号的一致性。对 ξ 进行进一步的分析发现，在接收区间不包含置信区间的情况下， ξ 的大小不仅和 $[a_i, b_i]$ 与 $[a, b]$ 的重叠部分的大小直接有关，而且还和它们的长度是否匹配有关。因此，应当保证置信区间与接收区间的长度匹配。

2 信号仿真的决策分析模型

2.1 置信区间和接收区间的长度匹配问题

置信区间与接收区间的长度是否匹配影响着 ξ 取值的大小。如果置信区间和接收区间的重叠程度很高，但长度不够匹配的话，例如置信区间比接收区间的长度大很多，此时随机数 ξ 的值会很小，这样就降低了 ξ 的取值变化的敏感性，不利于定量分析仿真信号的一致性。下面，研究置信区间与接收区间的长度匹配问题。

参数置信区间估计是受总体的样本容量影响的，增大或减小样本容量可以控制置信区间的长度^[8]。而置信区间与接收区间的长度匹配，可以理解为让它们的半长之差的绝对值很小。故可事先给定一个允许值，并以赋初值的形式给出第一次仿真实验的样本容量，就可以确定样本容量的一个范围。由于该范围是在给定接收区间的情况下求出的，它和接收区间之间存在的制约关系能够保证找到一个最合适的样本容量值，称为最佳样本容量，根据最佳样本容量值对信号仿真数据进行抽样，能够得到与接收区间长度匹配的置信区间。求最佳样本容量的具体步骤如下：

1) 给定信号仿真模型的某个特征参数 θ 的均

值 μ 的接收区间 $[a, b]$ ，置信区间与接收区间半长之差的绝对值的允许值 ε ，参数区间估计的置信水平 $1 - \alpha$ ；

2) 由给定接收区间得到接收区间的半长 $\frac{L}{2} = \frac{b-a}{2}$ ；第一次仿真实验对仿真系统抽取 N 个样本，由此得到特征参数 θ 的样本观测值 $x_{11}, x_{12}, \dots, x_{1N}$ ，进而得到置信区间的半长为 $\frac{\sigma}{\sqrt{n}} u_{1-\frac{\alpha}{2}}$ ，其中， σ 为特征参数 θ 的均方差，在 σ 未知时以其无偏估计量样本均方差 S^* 代替；

3) 取 σ 无偏估计量 S^* ，为了实现区间长度匹配，要求有式 (3) 成立。由式 (3) 得到样本容量值 n 范围，有式 (4)；

4) 由式 (4)，通过枚举法找出最佳样本容量值 Q 。

$$\left| \frac{S^*}{\sqrt{n}} u_{1-\frac{\alpha}{2}} - \frac{L}{2} \right| \leq \varepsilon \quad (3)$$

$$\left[\frac{u_{1-\frac{\alpha}{2}}^2 S^{*2}}{(\frac{L}{2} + \varepsilon)^2}, \frac{u_{1-\frac{\alpha}{2}}^2 S^{*2}}{(\frac{L}{2} - \varepsilon)^2} \right] \quad (4)$$

2.2 置信区间偏离函数

仿真信号模型一般都比较复杂，因此仿真试验往往要经过反复的调试才能达到令人满意的程度，不可能一次成功。故在对信号仿真模型进行阶段性评估后，在不能达到要求时如何更好、更快的进行调试，使调试的策略及步骤更加具有针对性、有效性，受到普遍关注。文献[9]给出了反映置信区间与接收区间位置关系的随机数 η ，对上述问题进行了讨论，并联合随机数 ξ 对 $\delta = \xi\eta$ 的数值情况，采用能量聚点方法进行分析，这样能较快地发现信号仿真数据的趋向性，为找到与实际信号序列的差异，提高调试效率提供了辅助决策。但不足之处是， δ 并没有充分地反映置信区间与接收区间的位置关系，原因是当 $\eta = 0$ 不能区分是区间中点对齐还是不相交，并且 η 没有充分反映出置信区间对接收区间的偏离程度。如图 1， $[a, b]$ 为接收区间， A 为该区间的中点， $[a_1, b_1]$ 、 $[a_2, b_2]$ 为仿真抽样所得置信区间， $A_i (i=1, 2)$ 分别为其中点， A_1 、 A_2 同在 A 的右侧，但对 A 的偏离程度不同。

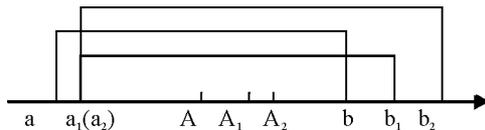


图 1 接收区间与置信区间的位置关系图

针对上述情况, 建立置信区间对于接收区间的偏离函数 $f(y)$ 来反映置信区间与接收区间的位置关系, 见式 (5)。在式 (5) 中, y 、 x_A 分别为置信区间、接收区间的中点的坐标。当 $y = x_A$ 时, 即 2 个区间的中点重合时, $f(y)=0$; 当 $y > x_A$ 时, 置信区间在接收区间右侧, y 对 x_A 的偏离越小, $f(y)$ 的值越从右侧接近于 0, y 对 x_A 的偏离越大, $f(y)$ 的值越接近于 1; 当 $y < x_A$ 时, 置信区间在接收区间左侧, y 对 x_A 偏离越小, $f(y)$ 的值越从左侧接近于 0, y 对 x_A 的偏离越大, $f(y)$ 的值越接近于 -1。可见, 偏离函数 $f(y)$ 相对于随机数 η 能更好地反映置信区间与接收区间的位置关系。

$$f(y) = \begin{cases} 1 - e^{-(y-x_A)} & \text{当 } y > x_A \text{ 时} \\ 0 & \text{当 } y = x_A \text{ 时} \\ -1 + e^{y-x_A} & \text{当 } y < x_A \text{ 时} \end{cases} \quad (5)$$

2.3 信号仿真决策分析模型

根据上述讨论, 建立了信号仿真决策分析模型, 见图 2。其具体步骤是:

- 1) 设定模型可信的评判标准: 信号特征参数 θ 的接收区间 $[a, b]$, 频率阈值 f_0 , 参数区间估计的最低置信度水平 $1-\alpha$;
- 2) 由给定的接收区间确定最佳样本容量值 Q (见 2.1 节);
- 3) 根据最佳样本容量值 Q , 对仿真输出信号抽取样本, 得到 N 组样本观测值, 由此计算得到的 N 个置信区间 $[a_i, b_i]$, 其中的置信度 $1-\alpha_i \geq 1-\alpha_0$, ($i=1, 2, \dots, N$);
- 4) 统计事件 $[a_i, b_i] \subseteq [a, b]$ 发生的频率 $f_N(A)$;
- 5) 对于 f_0 , 若 $f_N(A) \geq f_0$, 则在该评判标准下可认定信号仿真模型可信; 若 $f_N(A) < f_0$, 根据式 (2) 考察随机数 ξ , 并对 ξ ($i=1, 2, \dots, N$) 做统计分析, 如估计其均值、均方差等, 其结果可作为进一步调试模型的辅助决策参考;
- 6) 当 ξ 的均值的估计值 $\hat{\mu}$ 较小而不能满足要求时, 根据仿真抽样的数据计算偏离函数, 由此分析置信区间的偏离趋势, 并建立基于偏离函数的能

量聚点特征量^[10], 根据分析的结果对模型进行调试, 调试后重复步骤 2), 直至接受模型。

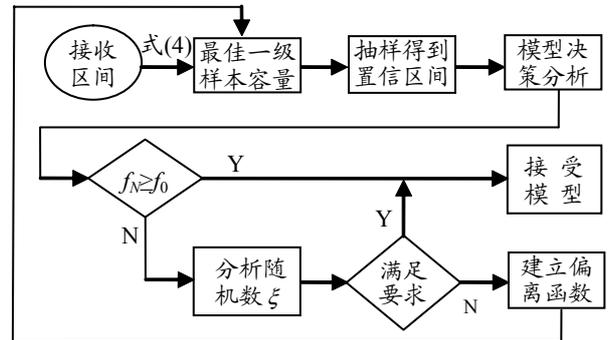


图 2 基于置信区间的信号仿真决策分析模型

3 实例分析

在水下航行体辐射噪声仿真模型中, 线谱基频率是重要的特征参数之一, 该特征参数是否满足要求是评估仿真信号能否逼真复现实体信号的重要标准。目标环境及相关数据为: 环境噪声级约 $\times \times$ dB; 工况为匀速、等深、直航, 航速 $\times \times$ 节, 航深 $\times \times$ 米; 分析频率范围 $\times \times$ Hz, 采样频率 40 kHz; 采样 4 096 点, 加 Hanning 窗。

3.1 最佳样本容量下的匹配置信区间

获取仿真输出的抽样信号 $x(t)$ 后, 采用时域平均的方法减少随机因素的影响, 并提取特征参数 θ 。给定线谱基频率均值的接收区间为 $[18.25, 18.75]$, 仿真抽样所得置信区间与接收区间半长之差的绝对值的允许值 $\epsilon = 0.125$, 置信水平 $1-\alpha = 0.95$ 。按 2.1 节求最佳一级样本容量的步骤, 先取出 40 个实验样本观测值: 18.0 18.6 18.8 18.2 18.6 18.2 18.3 18.6 18.5 18.1 18.5 18.8 18.5 18.8 18.7 18.5 18.4 18.5 18.8 18.3 18.6 18.2 18.6 18.5 18.2 18.3 18.5 18.2 18.5 18.5 18.3 18.4 18.0 18.5 18.1 18.4 18.4 18.8 18.2 18.4。由式 (4) 得样本容量值的范围是 $[1.32, 11.89]$, 并通过枚举法取最佳样本容量值 $Q=11$ 。按照 $Q=11$ 对仿真数据进行抽样, 得到的其中一个置信区间为 $[18.44, 18.92]$, 置信区间长度 $l=0.48$, 随机数 $\xi=0.85$ 。可见, 由最佳样本容量值确定的置信区间的长度和给定的接收区间的长度匹配良好。

3.2 基于偏离函数值的仿真模型调试

给定接收区间为 $[18.25, 18.75]$ 后, 对信号仿真模型的输出数据按照最佳样本容量进行抽样, 得到 30 个置信区间 $[a_i, b_i]$ 。经计算发现随机数 ξ 的值普遍偏低, 直观说明仿真信号一致性偏低, 因此建立偏离

函数 $f(y)$ ，对置信区间的偏离度进行定量考察。根据置信区间对接收区间的偏离度的实际情况，考察 $f(y)$ 在区间 $(-0.2, 0.2)$ 中的分布，取等长度为 0.05、相邻子区间依序重叠 50% 的滑动子区间 $(a, a+0.05)$ ，各子区间的 $f(y)$ 的值的散点分布情况，如图 3。由各子区间所含点数 M_n 求得的 $f(y)$ 的密度序列计算出密度序列 $\{\rho_n\}$ 的 λ 水平能量聚点，如图 4 ($\rho_n = M_n / 30$, $-0.2 \leq a \leq 0.15, n=1, 2, \dots, 30$)。

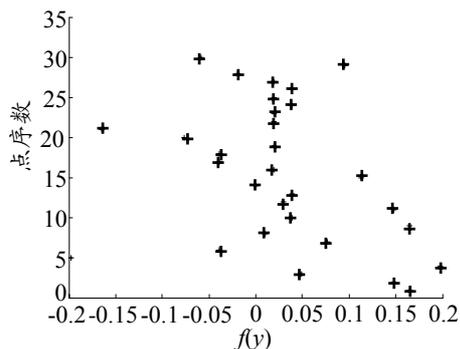


图 3 偏离函数 $f(y)$ 的值的散点图

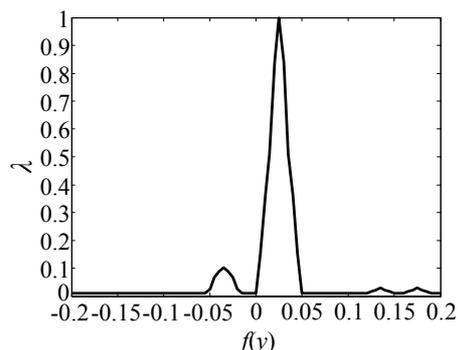


图 4 偏离函数的散点序列的能量聚点

由图 3 可以看出， $f(y)$ 值分布趋向在 0 至 0.05 之间，且总趋势偏右。根据 $\{\rho_n\}$ 的 λ 水平能量聚点的分析结果，针对性地对信号仿真模型进行调试，较快地提高了仿真信号的一致性。对比计算表明：与仅凭直觉经验进行调试的方法相比，根据偏离函数值分布进行调试能使调试次数平均减少 25.6%。

3.3 计算结果对比及结论

如文献[9]所述，利用随机数 ξ 对仿真抽样数据进行分析，能进一步挖掘输出信号数据当中的有用信息。下面，仅就解决置信区间和接收区间的匹配问题前后的差异进行对比。在不考虑置信区间和接收区间的匹配关系时，经计算得到的随机数 ξ 的均值的估计值为 0.61，而在考虑匹配问题时，得到的随机数 ξ 的均值的估计值为 0.78，随机数 ξ 的信息

分析效能有较大提高。可见，充分挖掘仿真信号的有用信息是对信号仿真模型准确评估的重要条件。

利用偏离函数对信号仿真模型进行有针对性的调试，再经计算得到 ξ 值为 0.83 0.83 0.83 0.87 0.80 0.84 0.79 0.81 0.81 0.85 0.88 0.86 0.85 0.86 0.82 0.84 0.78 0.84 0.85 0.86 0.87 0.83 0.84 0.82 0.81 0.79 0.88 0.85 0.84 0.85 0.81, ξ 的均值、均方差分别为 0.835 2、0.026 4。据此得出结论：辐射噪声的仿真信号的一致性估计为 0.835，仿真信号复现实际信号的逼真程度较高，能够满足工程需求。

4 结束语

该模型考虑了给定接收区间后仿真信号特征参数的置信区间与接收区间的匹配问题，有助于凸显仿真抽样数据所蕴涵的有用信息；同时，建立了置信区间对接收区间的偏离函数，分析特征参数的分布趋向，并利用分析后的反馈信息对信号仿真模型进行了修正，提高了模型的调试效率。但该模型仅讨论了一个特征参数的信号仿真决策分析，而仿真信号的特征参数一般不止一个，下一步，将对多参数的信号仿真决策分析模型进行研究。

参考文献:

- [1] 魏华梁, 王肇敏, 刘藻珍. 系统仿真置信度研究中的若干问题与准则[J]. 系统仿真学报, 2000, 12(1): 39-42.
- [2] 孙勇成, 马艳琴, 周献中, 等. 建模与仿真中验证与确认的测试与评估实现[J]. 系统仿真学报, 2005, 17 (10): 2317-2320.
- [3] 李鹏. 舰艇鱼雷初期弹道仿真可信度的相似评估[J]. 兵工自动化, 2008, 27 (7): 18-20.
- [4] 胡鹏, 邵晨曦, 杨明, 等. 一种军事仿真概念模型的完整性评估方法[J]. 系统仿真学报, 2008, 20(16): 4207-4210.
- [5] 朱秋煜, 王朔中. 图像特征检测和马氏距离中的数据融合与置信度[J]. 电子与信息学报, 2008, 03(30): 534-538.
- [6] 焦鹏, 唐见兵, 查亚兵. 仿真可信度评估中相似度方法的改进及其应用[J]. 系统仿真学报, 2007, 19(12): 2658-2660.
- [7] 邢继娟, 葛含益, 卜广志. 仿真模型置信度分析方法研究[J]. 军事运筹与系统工程, 2005, 19(4): 35-39.
- [8] 庄楚强, 何春雄. 应用数理统计基础[M]. 第 3 版. 广州: 华南理工大学出版社, 2006.
- [9] 胡伟文, 苑秉成, 杨鹏. 基于置信区间的仿真试验结果决策分析方法[J]. 系统工程与电子技术, 2009, 03(31): 621-624.
- [10] 胡伟文, 姜礼平, 苑秉成. 能量聚点与特征提取方法研究[J]. 数据采集与处理, 2004, 19(4): 417-420.