



UNIVERSIDADE FEDERAL DE PERNAMBUCO

**DEPARTAMENTO DE ENGENHARIA DE PRODUÇÃO
CURSO DE GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO**

**INFERÊNCIA BAYESIANA PARA AVALIAÇÃO DE
MÉTRICAS DE CONFIABILIDADE**

MÁRCIO JOSÉ DAS CHAGAS MOURA

Orientador: Prof. Enrique López Droguett, Ph. D.

Recife, abril / 2005

M929i

Moura, Márcio José das Chagas

Inferência bayesiana para avaliação de métricas de confiabilidade / Márcio José das Chagas Moura. – Recife: O Autor, 2005.

ix, 39 f.; il., gráfs., tabs.

Monografia (TCC) – Universidade Federal de Pernambuco. CTG. Depto. de Engenharia de Produção, 2005.

Inclui referências bibliográficas.

1. Engenharia de Produção. 2. Confiabilidade. 3. Teorema de Bayes. 4. Markov Chain Monte Carlo - MCMCI. Título.

658.5 CDD (22.ed.)

UFPE/BCTG/2006-115

AGRADECIMENTOS

Esse sem dúvida é um dos tópicos mais fácil de redigir. As palavras fluem naturalmente. A causa disso é o sentimento de alegria e alívio desse momento. Inúmeras pessoas contribuíram para a concretização desse sonho. Em primeiro lugar, agradeço a Deus por te me concedido as pessoas mais importantes da minha vida. O ditado diz que atrás de um grande homem existe sempre uma mulher. A vida me deu a dádiva de ter não apenas uma, mas três. Minha avó, Iracema das Chagas (em memória), minha mãe, M.^a da Conceição e minha filha Geovanna. São essas três mulheres que fazem minha vida ter sentido a cada dia que acordo. Agradeço à prof.^a Dayse Duarte e ao prof. Enrique López por terem acreditado em mim e me dado à oportunidade de trabalhar no Riscotec – Grupo de Pesquisas em Análise de Riscos e Eng^a de Confiabilidade. Agradeço também a todos os meus amigos da graduação em Eng^a de Produção e a todos os indivíduos que compõem o Departamento de Eng^a de Produção.

SUMÁRIO

RESUMO	VIII	
ABSTRACT	IX	
1	INTRODUÇÃO	1
1.1	JUSTIFICATIVA	2
1.2	OBJETIVOS	2
1.2.1	<i>Objetivo Geral</i>	2
1.2.2	<i>Objetivos Específicos</i>	2
2	TEOREMA DE BAYES.....	4
2.1	CONCEITOS BÁSICOS.....	4
2.2	INFERÊNCIA BAYESIANA	5
2.2.1	<i>Introdução</i>	5
2.2.2	<i>Probabilidade Subjetiva</i>	6
2.2.3	<i>Teoria Amostral versus Inferência Bayesiana</i>	7
2.2.3.1	Inferência Clássica	7
2.2.3.2	Inferência Bayesiana	8
3	A METODOLOGIA MARKOV CHAIN MONTE CARLO (MCMC).....	11
3.1	INTEGRAÇÃO MONTE CARLO.....	12
3.2	INTRODUÇÃO A CADEIAS DE MARKOV	12
3.3	METODOLOGIA DE SIMULAÇÃO MCMC.....	14
3.4	O MÉTODO DE ACEITAÇÃO – REJEIÇÃO (A-R).....	16
3.5	O ALGORITMO METROPOLIS – HASTINGS (M – H).....	18
3.6	ESCOLHA DA DISTRIBUIÇÃO CANDIDATA $Q(X, Y)$	21
4	ESTUDO DE CASO	24
4.1	ELICITAÇÃO DA OPINIÃO DO ESPECIALISTA	24
4.1.1	<i>Métodos de Elicitação da Opinião do Especialista</i>	24
4.2	FUNÇÃO DE VEROSSIMILHANÇA.....	27
4.3	DISTRIBUIÇÃO A PRIORI	29
4.4	APROXIMAÇÃO DA DISTRIBUIÇÃO A POSTERIORI.....	31
5	RESULTADOS E DISCUSSÃO.....	32
6	CONCLUSÕES	36
6.1	DESAFIOS FUTUROS	37
REFERÊNCIAS BIBLIOGRÁFICAS.....	38	

LISTA DE FIGURAS

FIGURA 2-1 – INFERÊNCIA BASEADA NA TEORIA AMOSTRAL	8
FIGURA 3-1 – MÉTODO ACEITAÇÃO – REJEIÇÃO.....	17
FIGURA 3-2 – CALCULANDO PROBABILIDADES DE DESLOCAMENTO COM FUNÇÃO GERADORA DE CANDIDATAS SIMÉTRICA	20
FIGURA 4-1 – ELICITANDO A OPINIÃO DO ESPECIALISTA	31
FIGURA 5-1 – VALORES GERADOS PARA λ PELO ALGORITMO M-H	34
FIGURA 5-2 – FUNÇÃO DE DISTRIBUIÇÃO ACUMULADA PARA λ (CDF).....	34
FIGURA 5-3 – FUNÇÃO DE DENSIDADE DE PROBABILIDADE PARA λ (PDF).....	34

LISTA DE TABELAS

TABELA 4-1 – DINHEIRO REQUERIDO PARA DUPLICAR A FELICIDADE.....	26
TABELA 5-1 – DADOS DE FALHA EMPÍRICOS	32
TABELA 5-2 – ELICITAÇÃO DA OPINIÃO DO ESPECIALISTA	33
TABELA 5-3 – PARÂMETROS DA DISTRIBUIÇÃO BETA.....	33

LISTA DE EQUAÇÕES

EQUAÇÃO 2-1.....	4
EQUAÇÃO 2-2	5
EQUAÇÃO 2-3	7
EQUAÇÃO 2-4	8
EQUAÇÃO 2-5	8
EQUAÇÃO 2-6.....	8
EQUAÇÃO 2-7.....	9
EQUAÇÃO 2-8.....	9
EQUAÇÃO 2-9.....	9
EQUAÇÃO 2-10.....	9
EQUAÇÃO 2-11	9
EQUAÇÃO 2-12	9
EQUAÇÃO 2-13.....	10
EQUAÇÃO 2-14	10
EQUAÇÃO 2-15	10
EQUAÇÃO 2-16.....	10
EQUAÇÃO 3-1.....	11
EQUAÇÃO 3-2.....	11
EQUAÇÃO 3-3.....	12
EQUAÇÃO 3-4.....	13
EQUAÇÃO 3-5.....	13
EQUAÇÃO 3-6.....	13
EQUAÇÃO 3-7.....	14
EQUAÇÃO 3-8.....	14
EQUAÇÃO 3-9.....	14
EQUAÇÃO 3-10.....	15
EQUAÇÃO 3-11.....	16
EQUAÇÃO 3-12.....	16
EQUAÇÃO 3-13	18
EQUAÇÃO 3-14.....	18
EQUAÇÃO 3-15.....	19
EQUAÇÃO 3-16.....	19
EQUAÇÃO 3-17.....	19
EQUAÇÃO 3-18.....	21
EQUAÇÃO 3-19.....	22
EQUAÇÃO 4-1.....	25
EQUAÇÃO 4-2.....	28
EQUAÇÃO 4-3.....	28
EQUAÇÃO 4-4.....	29
EQUAÇÃO 4-5.....	29
EQUAÇÃO 4-6.....	30
EQUAÇÃO 5-1.....	32

RESUMO

Consumidores esperam que os equipamentos adquiridos funcionem como anunciado pelo fabricante. Assim, o objetivo do fabricante é desenvolver e comercializar produtos com custos reduzidos e em curtos espaços de tempo que satisfaçam as necessidades dos consumidores. Pode – se dizer que a confiabilidade é um atributo inerente a um produto o qual deve ser considerado desde o processo de concepção do produto até a chegada do mesmo às mãos do consumidor. Porém, a escassez de dados é uma característica cotidianamente encontrada quando se deseja avaliar a confiabilidade de determinado equipamento. Esta situação é acarretada por diversas razões, tais como dificuldades em obter dados de falha consideráveis para o produto sob condições normais de uso, pressões para alcançar prazos e limitações de custo. Desta forma, torna-se determinante a utilização de fontes de informação alternativas, tais como testes de vida acelerados e opiniões de especialistas para a avaliação da confiabilidade de um equipamento. Entenda por especialistas, engenheiros ou outro indivíduo ou grupo de indivíduos com treinamento adequado e familiarizado com o equipamento e processos que constituem o objeto da análise. Apresentar-se-á aqui uma metodologia para estimativa de métricas de confiabilidade baseada no Teorema de Bayes e na metodologia de simulação Markov Chain Monte Carlo (MCMC). Um Estudo de Caso será apresentado para uma Bomba de resfriamento utilizada em uma Usina Nuclear.

Palavras – chaves: Confiabilidade, Teorema de Bayes, MCMC

ABSTRACT

Consumers hope that the acquired equipment functions as announced for the manufacturer. Thus, the objective of the manufacturer is to develop and to commercialize products with reduced costs and in short spaces of time that satisfy the necessities of the consumers. It can say that the Reliability is an inherent attribute to a product which must be considered since the process of conception of the product until the arrival of the same one to the hands of the consumer. However, the scarcity of data is a characteristic daily found when it is desired to evaluate the Reliability of given equipment. This situation is caused by diverse reasons, such as difficulties in getting considerable failure data for the product under normal conditions of use, pressures to reach stated periods and limitations of cost. Of this form, the use of alternative sources of information becomes determinative, such as accelerated life testing and opinions of specialists for the evaluation of the Reliability of equipment. It understands for specialists, engineers or another individual or group of individuals with adequate training and made familiar to the equipment and processes that constitute the object of the analysis. A methodology for estimate of metric of Reliability will be presented here based in the Theorem of Bayes and the methodology of simulation Markov Chain Monte Carlo (MCMC). A Study of Case it will be presented for a Pump of cooling used in a Nuclear Plant.

Key – Words: Reliability, Theorem of Bayes, MCMC

1 INTRODUÇÃO

A análise de confiabilidade é uma técnica de suporte a tomadas de decisão e controle que auxilia gestores na busca da garantia da execução satisfatória das funções dos itens a respeito de um dado sistema, considerando suas limitações, o seu desgaste e os fatores que influenciam seu desempenho, sejam estes itens equipamentos ou pessoas (Firmino, 2004).

Em um mercado competitivo como o atual, a busca permanente por vantagens competitivas é imperativa para sobrevivência da empresa. Dentre *os características da qualidade*, i.e., elementos de decisão que um consumidor elege para adquirir determinado produto (Paladini, 2000), a confiabilidade aparece entre os mais bem ranqueados ao lado da manutenibilidade e da qualidade.

A análise da confiabilidade de um equipamento proporciona redução de custos, otimização do processo produtivo, estabelecimento de intervalos mais acurados de manutenção preventiva, redução do tempo de espera em operações de *Setup*, níveis mais baixos de estoque intermediários, entre outras vantagens.

Basicamente, existem dois modelos para análise de confiabilidade, quais sejam: o modelo Clássico e o modelo Bayesiano. Através do modelo clássico, para avaliar a confiabilidade de determinado equipamento se faz necessário grandes amostras de falha nas condições de uso nominais. Entretanto, esses dados são difíceis de serem obtidos devido a causas, tais como: elevado tempo de vida útil do produto, período curto entre projeto e lançamento do produto no mercado, dificuldade em testar produtos de uso contínuo sob condições de uso normais. Portanto, é muito comum para engenheiros no dia a dia de seu trabalho se deparar com situações onde a escassez de dados é um entrave determinante para avaliação da confiabilidade de determinado equipamento.

Através do modelo Bayesiano para avaliação da confiabilidade é permitido que fontes alternativas de falha sejam utilizadas, quando dados de campo ou de teste de equipamentos estão indisponíveis. Dentre essas fontes alternativas, surge como alavancador da Teoria Bayesiana a denominada *Opinião do Especialista*, onde é possível que o conhecimento implícito de um indivíduo inteirado com o objeto de análise seja externalizado através de uma distribuição de probabilidade, a qual denominamos de distribuição a priori. Além disso, à medida que novas informações sobre o objeto de estudo vão surgindo, retroalimentações do conhecimento sobre tal são realizadas, consistindo conseqüentemente num processo de melhoria contínua.

1.1 Justificativa

Na tentativa de desvendar e quantificar métricas de confiabilidade, engenheiros e profissionais da área são frequentemente deparados com situações caracterizadas pela escassez de dados relevantes para essa avaliação.

As dificuldades encontradas na utilização de métodos estatísticos tradicionais, baseados na análise de dados através da ótica frequentista, têm sido em grande parte os responsáveis pelo grande sucesso e popularidade da avaliação da confiabilidade através do método Bayesiano. A base para este procedimento se origina a partir do trabalho do reverendo Bayes publicado em 1763, dois anos após a sua morte. O método tem se mostrado extremamente útil em confiabilidade e em todas as áreas de aplicação estatística.

Tal método permite que várias fontes de informação relativas ao problema sejam conjugadas incrementando o conhecimento prévio sobre algo. Dentre tais fontes de informação, destaca-se a opinião de especialistas, na qual é possível que o know-how de um indivíduo ou grupo de indivíduos devidamente ambientado com o problema seja utilizado para a quantificação de determinados atributos de confiabilidade.

1.2 Objetivos

Este trabalho apresenta um objetivo geral e alguns objetivos específicos que são descritos a seguir.

1.2.1 Objetivo Geral

Deseja-se com este trabalho aplicar o Teorema de Bayes e a metodologia Markov Chain Monte Carlo (MCMC) no problema de inferência Bayesiana, enfatizando para utilidade do algoritmo de simulação Metropolis-Hastings (M-H).

1.2.2 Objetivos Específicos

A estratégia adotada para a realização do trabalho percorre os três temas aqui enfatizados: Teorema de Bayes, a metodologia MCMC e o algoritmo M-H. Deseja-se mostrar como a metodologia MCMC e algoritmo M-H são utilizados dentro do contexto do problema da inferência Bayesiana.

Será apresentado um apanhado geral sobre todos os temas aqui abordados, constituindo a fundamentação teórica deste trabalho. Discutir-se-á o problema da inferência Bayesiana, assim como assuntos relacionados à metodologia MCMC, ao algoritmo M-H e suas respectivas possíveis aplicações. Um caso de aplicação será apresentado a fim de mostrar como o Teorema de Bayes permite a adaptação de informações vindas de fontes distintas, como opiniões de especialistas e/ou dados empíricos, permitindo assim atualizações sucessivas do conhecimento sobre algum parâmetro de interesse.

Pode –se enunciar, portanto, os seguintes objetivos específicos:

- Analisar via revisão bibliográfica o problema da inferência Bayesiana e a metodologia Markov Chain Monte Carlo (MCMC);
- Analisar a aplicabilidade de métodos MCMC no problema de Inferência Bayesiana, especialmente o algoritmo Metropolis – Hastings;
- Apresentar um Estudo de Caso para Bombas que compõem um sistema de resfriamento de uma Usina Nuclear de forma a avaliar métricas de confiabilidade através da aplicação conjunta do Teorema de Bayes e da metodologia MCMC.

2 TEOREMA DE BAYES

2.1 Conceitos Básicos

Através da aplicação da análise Bayesiana em confiabilidade é possível para o engenheiro usar informações provenientes de fontes distintas, como handbooks e opiniões de especialistas, na formulação de uma função de probabilidade para algum parâmetro de interesse. Para um dado parâmetro, por exemplo, a taxa de falha, os resultados podem ser sistematicamente atualizados à medida que novas informações tornam-se disponíveis.

Uma simples visão do Teorema de Bayes é apresentada através da teoria de probabilidade. Considere que estamos interessados em um determinado evento A , como por exemplo, o nível de confiabilidade a ser atingido por um equipamento, e que E represente alguma nova informação relevante à avaliação de A , como o resultado de um teste acelerado. Então, o Teorema de Bayes estabelece que a probabilidade do evento A dado a nova evidência (dados) E é proporcional ao produto entre a probabilidade do evento A antes de obtermos a nova informação E , e a probabilidade de observar a evidência E caso o evento A ocorresse.

Ou seja,

$$P(A|E) = \frac{P(E|A)P(A)}{P(E)}$$

Equação 2-1

, onde $P(A)$ é conhecida como a *probabilidade a priori de A*, i.e., antes de tomarmos conhecimento de E ; $P(E|A)$ é a probabilidade de que a evidência E seja observada se A é realmente verdadeiro (ocorre), e é conhecida como *função de verossimilhança*; e $P(A|E)$ é a probabilidade a posteriori de A , i.e., após termos obtido a nova informação representada por E . Logo, $P(A|E)$ representa a nossa probabilidade atualizada sobre o evento A uma vez que obtemos a informação adicional E relevante a A .

Em geral, entretanto, estamos interessados na avaliação de atributos de confiabilidade ou parâmetros contínuos, como a taxa de falha. Na forma contínua, o Teorema de Bayes fica:

$$\pi(\theta | E) = \frac{L(E | \theta)\pi_o(\theta)}{\int_{\theta} L(E | \theta)\pi_o(\theta)d\theta}$$

Equação 2-2

, onde θ representa algum parâmetro de interesse, como a taxa de falha ou de reparo de um equipamento. Como anteriormente, $\pi_o(\theta)$ é a distribuição a priori sobre θ representando a totalidade da informação disponível ao engenheiro sobre o parâmetro de interesse antes de observar a nova evidência E (como resultados de testes); $L(E | \theta)$ é análoga a $P(E | A)$, conhecida como *função de verossimilhança*. Note que a informação representada por E convencionalmente está na forma de dados empíricos (como testes). Outras fontes de informação também podem ser utilizadas, como por exemplo, a opinião de especialistas e dados não completamente relevantes, como dados de falha de um equipamento semelhante ao do objeto de estudo, mas que opera em condições ambientais distintas.

O Teorema de Bayes é uma ferramenta ideal para realizarmos atualizações sucessivas do que nós, engenheiros, sabemos sobre a confiabilidade (ou qualquer outro parâmetro de interesse) de determinado equipamento à medida que novas informações se tornam disponíveis, funcionando como um mecanismo de retroalimentação do conhecimento. Basicamente, a distribuição a posteriori de hoje torna-se a distribuição a priori de amanhã. O processo de atualização pode ser expresso da seguinte forma:

$$priori + dados(evidência) \Rightarrow posteriori(\pi(\theta | E))$$

2.2 INFERÊNCIA BAYESIANA

2.2.1 Introdução

Rev. Thomas Bayes (1702-1761), em seu famoso artigo publicado postumamente, *An Essay Towards Solving a Problem in the Doctrine of Chances*, fornece a base para o método conhecido como “Inferência Estatística Bayesiana”. Nele, Bayes combina dois corpos de evidência com o objetivo de realizar um refinamento da informação. É possível combinar evidência epistêmica (probabilidade *a priori*; opinião do especialista, etc.) com frequentista (dados, frequência relativa, função de verossimilhança, etc.). Esta combinação gera uma probabilidade dita *a posteriori*.

Devido à importância do artigo, o mesmo tem sido bastante republicado. Desde a sua publicação, o interesse tem alternado entre períodos de aceitação e rejeição do método, como base para a inferência estatística (Martz & Waller, 1982).

Segundo Martz & Waller (1982), alguns fatores contribuem para o recente ressurgimento.

Primeiro, o trabalho de alguns autores, tais como DeFinetti (1937), Good (1950), Jeffreys (1961), Lindley (1965), Ramsey (1931), e Savage (1954), tem fornecido a base filosófica para este método.

O segundo argumento diz que as outras teorias se baseiam em pré-consideração restritiva, que conduzem a análise muitas vezes a algo fora da realidade, estas considerações são basicamente para fornecer uma solução matemática conveniente.

O terceiro se baseia no fato do surgimento de computadores de alta velocidade, permitindo que uma vasta classe de análise de dados possa ser tanto considerada quanto comparada.

O quarto argumento está relacionado ao interesse dos especialistas por métodos que incorporem e combinem vários recursos de dados, i.e., tanto dados subjetivos quanto objetivos para uma análise.

2.2.2 Probabilidade Subjetiva

A probabilidade subjetiva não trata somente de *eventos*, mas também de *proposições*. Nela, a probabilidade reflete o nível de confiança em uma *proposição* (Martz & Waller, 1982). Uma *proposição* é considerada como uma coleção de eventos, que não pode ser interpretada como uma série de repetições (i.e., na prática a probabilidade dita subjetiva não pode ser adquirida por meio de experimentos, tal como um jogo de dados, desta forma é dito que a probabilidade subjetiva trabalha com proposições). Uma proposição tem a característica de poder ser provada verdadeira ou falsa. No jargão estatístico a palavra *hipótese* é usada no lugar da palavra *proposição*, isto se dá pelo fato da hipótese está relacionada ao interesse sobre a proposição aceita como verdadeira. Uma nova evidência eleva a relevância da *hipótese* e então muda o grau de crença sobre ela.

Segundo Martz & Waller (1982), a probabilidade subjetiva reflete o grau de crença a respeito de uma *proposição*.

O grau de confiança (crença) é uma medida do conhecimento a respeito de uma proposição ou evento. Por conta disto ele é subjetivo por essência. É uma medida de incerteza e que representa o estado mental e não o mundo externo. (Nota de aula de Droguett, 2002).

2.2.3 Teoria Amostral versus Inferência Bayesiana

A inferência baseada na teoria amostral é diferente da realizada na inferência bayesiana. São duas escolas diferentes com relação a interpretação e consequentemente aplicação da probabilidade, dita frequentista (clássica) e subjetivista (Bayesiana).

2.2.3.1 Inferência Clássica

Na inferência clássica, a probabilidade é o limite da frequência relativa da ocorrência de um resultado desejado pelo número de vezes em que o experimento em questão é realizado. Formalmente, a probabilidade de um evento E é dado por:

$$\Pr(E) = \lim_{N \rightarrow \infty} \frac{N_E}{N}$$

Equação 2-3

, onde N_E é o número de vezes em que o evento desejado (E) ocorre;

N é o número total de tentativas (repetições);

$\Pr(E)$ é a razão se tal limite existe.

Significa que, a probabilidade de ocorrência do Evento E é igual à relação entre o número de vezes que o evento desejado (E) ocorre pelo número total de tentativas (repetições), quando o número total de tentativas tende ao infinito.

De acordo com a **Figura 2-1**, o método da teoria amostral tem início na postulação do modelo amostral mais adequado para os dados existentes, em seguida um argumento indutivo é usado na conjunção da observação amostral produzindo inferência acerca do parâmetro do modelo assumido.

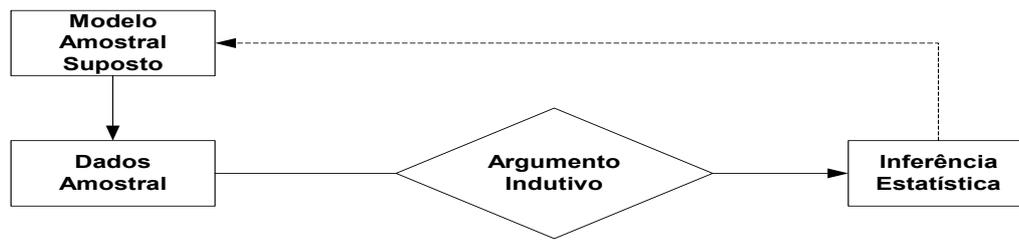


Figura 2-1 – Inferência baseada na teoria amostral

2.2.3.2 Inferência Bayesiana

Para um melhor entendimento de como a inferência Bayesiana opera, um exemplo será dado:

Exemplo 1: *A avaliação inicial de um engenheiro sobre os possíveis valores da taxa de falha λ de um componente, é representada na forma da distribuição de probabilidade que segue:*

$$Pr_o(\lambda) = (2000)^2 \lambda e^{-2000\lambda}, 0 \leq \lambda \leq \infty$$

Equação 2-4

O componente opera por 10.000 horas durante as quais o mesmo falha uma vez. O engenheiro em questão gostaria de obter a distribuição de probabilidade atualizada de λ dado a nova informação. A evidência dada no problema diz que $E = \{1 \text{ falha em } 10.000 \text{ horas}\}$.

Diante desta nova informação, o Teorema de Bayes é utilizado para atualizar o estado de conhecimento sobre a taxa de falha do componente.

$$Pr(\lambda / E) = \frac{L(E / \lambda) Pr_o(\lambda)}{L(E)}$$

Equação 2-5

Para modelar a evidência fornecida, será utilizada a distribuição de Poisson, uma vez que nela a evidência possui uma relação entre o número de observações desejado em função do tempo. Logo, a função de verossimilhança será dada por:

$$L(1 \text{ falha em } 10.000 / \lambda) = (10.000) \lambda e^{-10000\lambda}$$

Equação 2-6

A partir desta função de verossimilhança e da distribuição a priori, o denominador $L(E)$ da Equação 2-5 pode ser calculado:

$$L(E) = \int_0^{\infty} L(E/\lambda) \Pr_o(\lambda) d\lambda$$

Equação 2-7

$$L(E) = \int_0^{\infty} (10.000)\lambda e^{-10.000\lambda} (2.000)^2 \lambda e^{-2.000\lambda} d\lambda$$

Equação 2-8

$$L(E) = \int_0^{\infty} (4 * 10^{10} \lambda^2 e^{-12.000\lambda}) d\lambda$$

Equação 2-9

$$L(E) = 4.53 * 10^{-2}$$

Equação 2-10

A Equação 2-10 representa um coeficiente normalizador que diz respeito à probabilidade total da evidência E ocorrer. Desta maneira, a distribuição a posteriori para a taxa de falha será determinada a partir da nova evidência E oriunda dos dados de falha do componente.

Antes de iniciar a dedução do Teorema de Bayes, um breve comentário sobre probabilidade condicional será realizado. Seja Ω um espaço amostral, A um evento e E uma evidência. O símbolo $P(A/E)$ é a *probabilidade condicional do evento A, uma vez que E ocorreu*. Quando calculamos $P(A/E)$ tudo se passa como se E fosse o novo espaço amostral “reduzido” dentro do qual, queremos calcular a probabilidade de A .

Dado o evento A e a evidência E , temos que:

$$\Pr(A \cap E) = \Pr(E) \Pr(A/E)$$

Equação 2-11

A Equação 2-11 diz que, a probabilidade da ocorrência simultânea de dois eventos ($P(A \cap B)$) é o produto da probabilidade de um deles pela probabilidade do outro, dado o primeiro, i.e., a probabilidade da ocorrência simultânea do evento A e a evidência E , é igual ao produto da probabilidade da evidência E pela probabilidade do evento A dado a ocorrência da evidência E .

$$\Pr(E \cap A) = \Pr(A) \Pr(E/A)$$

Equação 2-12

A Equação 2-12 segue a mesma definição mostrada na Equação 2-11.

$$\Pr(A \cap E) = \Pr(E \cap A)$$

Equação 2-13

A Equação 2-13 expressa a igualdade entre as intersecções, de acordo com a propriedade comutativa.

$$\Pr(A/E)\Pr(E) = \Pr(E/A)\Pr(A)$$

Equação 2-14

A Equação 2-14 mostra a substituição das parcelas da igualdade apresentadas na Equação 2-13.

$$\Pr(A/E) = \frac{\Pr(E/A)\Pr(A)}{\Pr(E)}$$

Equação 2-15

Na Equação 2-15, a parcela $\Pr(A/E)$ é isolada, e assim o Teorema de Bayes é obtido, onde:

$\Pr(A/E)$ é denominado a probabilidade *a posteriori* de A dada a evidência E , que significa a ocorrência da evidência E .

$\Pr(E/A)$ é a *verossimilhança* da evidência assumindo a ocorrência de A .

$\Pr(A)$ é a probabilidade *a priori* do evento A .

$\Pr(E)$ é o termo de normalização (i.e., tem a função de garantir que a probabilidade a posteriori esteja compreendida entre 0 e 1).

$$\Pr(E) = \Pr(E/A)\Pr(A) + \Pr(E/\bar{A})\Pr(\bar{A})$$

Equação 2-16

A Equação 2-16 diz que a probabilidade da evidência E é igual ao produto da probabilidade da ocorrência da evidência E dado a ocorrência do evento A pela probabilidade de ocorrência do evento A , somado ao produto da probabilidade da ocorrência da evidência E dado a NÃO ocorrência do evento A pela probabilidade da NÃO ocorrência do evento A .

3 A METODOLOGIA MARKOV CHAIN MONTE CARLO (MCMC)

É muito raro que integrais que aparecem em problemas reais de Inferência Bayesiana possam ser avaliadas em formas fechadas. Casos como o Exemplo 1, não são comuns na vida real. Este foi um dos motivos que não permitiu que o Teorema de Bayes fosse imediatamente difundido: a dificuldade em se avaliar distribuições a posteriori que não podiam ser tratadas analiticamente.

A maioria das aplicações da metodologia Markov Chain Monte Carlo (MCMC) são orientadas para inferência Bayesiana. De uma perspectiva Bayesiana, não há distinção fundamental entre observações e parâmetros de um modelo estatístico: todos são considerados quantidades aleatórias. Permita que D denote os dados observados e θ sejam os parâmetros do modelo. A inferência Bayesiana formal então exige uma distribuição de probabilidade conjunta $P(D, \theta)$ ao longo de todas as quantidades aleatórias. Esta distribuição conjunta compreende duas partes: uma distribuição a priori $P(\theta)$ e uma função de verossimilhança $P(D|\theta)$. Especificando $P(\theta)$ e $P(D|\theta)$ produz um modelo de probabilidade completo, no qual:

$$P(D, \theta) = P(D|\theta)P(\theta)$$

Equação 3-1

Tendo observado D , o Teorema de Bayes é usado para determinar a distribuição de θ condicional em D :

$$P(\theta|D) = \frac{P(\theta)P(D|\theta)}{\int P(\theta)P(D|\theta)d\theta}$$

Equação 3-2

Esta é denominada distribuição posteriori de θ , e é o objeto de estudo de toda inferência Bayesiana.

A metodologia MCMC fornece um enorme escopo para modelagem estatística. MCMC é essencialmente integração Monte Carlo usando Cadeias de Markov. Bayesianos e alguns frequentistas necessitam solucionar algumas integrais possivelmente ao longo de distribuições de probabilidade em altas dimensões a fim de fazer inferências sobre os parâmetros do

modelo ou fazer previsões. Bayesianos necessitam integrar ao longo de distribuições posteriores de parâmetros do modelo e frequentistas podem necessitar integrar ao longo da distribuição das observações dados os valores dos parâmetros do modelo.

3.1 INTEGRAÇÃO MONTE CARLO

A integração Monte Carlo amostra valores de uma distribuição requerida e então forma médias amostrais para aproximar esperanças (médias). Markov Chain Monte Carlo amostra estes valores ‘simulando’ habilmente uma Cadeia de Markov por um determinado período de tempo. Há muitos caminhos para construção dessas cadeias, mas todos eles, incluindo o amostrador de Gibbs, são casos especiais de um geral framework desenvolvido por Metropolis et al. (1953) e Hastings (1970).

A idéia básica da integração Monte Carlo é avaliação de integrais, definidas ou não, através da geração de variáveis aleatórias no intervalo de $[0, 1]$. Se U_1, \dots, U_k são variáveis aleatórias uniformes independentes, segue que $g(U_1), \dots, g(U_k)$ são independentes e identicamente distribuídas tendo média θ , onde $\theta = \int g(x)dx$. Portanto, pela *Lei dos Grandes Números* segue que,

$$\sum_{i=1}^k \frac{g(U_i)}{k} \rightarrow E[g(U)] = \theta, k \rightarrow \infty$$

Equação 3-3

Portanto, nós aproximamos θ gerando um grande número de variáveis aleatórias u_i e tomando como nossa aproximação o valor esperado (esperança) de $g(u_i)$. Este método de aproximação de integrais é denominado de *Monte Carlo* (para maiores detalhes, veja Ross, 2002).

3.2 INTRODUÇÃO A CADEIAS DE MARKOV

Antes de apresentar o algoritmo Metropolis – Hastings, alguns comentários introdutórios sobre Cadeias de Markov devem ser feitos. Permita X_t denotar o valor de uma variável aleatória no instante t e que o espaço amostral seja o conjunto de possíveis valores de X . A variável aleatória é um processo de Markov se as probabilidades de transição entre diferentes valores no espaço amostral dependem apenas do corrente estado da variável aleatória, i.e.,

$$\Pr(X_{t+1} = s_j | X_0 = s_k, \dots, X_t = s_i) = \Pr(X_{t+1} = s_j | X_t = s_i)$$

Equação 3-4

Assim, para uma variável aleatória Markoviana o estado futuro depende apenas condicionalmente do estado presente da variável aleatória. O conhecimento dos valores anteriores da variável aleatória não modifica a probabilidade de transição. Uma cadeia de Markov é uma seqüência de variáveis aleatórias (X_0, \dots, X_n) geradas por um processo de Markov. Uma cadeia de Markov particular é definida mais completamente por suas probabilidades de transição (ou Kernel de Transição), $P(i, j) = P(i \rightarrow j)$, que é a probabilidade que o processo desloque-se do estado i para o j , isto é:

$$P(i, j) = P(i \rightarrow j) = \Pr(X_{t+1} = s_j | X_t = s_i)$$

Equação 3-5

Freqüentemente se usa a notação $P(i \rightarrow j)$ para se referir um deslocamento de i para j , enquanto muitos textos definem $P(i, j) = P(j \rightarrow i)$, então usaremos a notação utilizando a seta para evitar confusão.

Permita que $\pi_j(t) = \Pr(X_t = s_j)$ denote a probabilidade que a cadeia esteja no estado j no instante de tempo t . Iniciamos a cadeia especificando um vetor inicial $\pi(0)$. Freqüentemente, todos os elementos de $\pi(0)$ são zero exceto por um elemento singular de valor 1, indicando que o processo começa em um particular estado. No contexto de confiabilidade este estado corresponde ao estado operacional do sistema.

A probabilidade que a cadeia esteja em um estado s_i em um instante de tempo $t+1$ é dada pela *equação de Chapman – Kolmogorov*:

$$\pi_i(t+1) = \Pr(X_{t+1} = s_i) = \sum_k \Pr(X_{t+1} = s_i | X_t = s_k) * \Pr(X_t = s_k) = \sum_k P(k \rightarrow i) \pi_k(t) = \sum_k P(k, i) \pi_k(t)$$

Equação 3-6

Evoluções sucessivas da *equação de Chapman – Kolmogorov* descrevem a evolução da cadeia.

Nós podemos escrever as equações de *Chapman – Kolmogorov* na forma matricial como segue. Defina a matriz P das probabilidades de transição como a matriz a qual o i, j –ésimo elemento é $P(i, j)$, a probabilidade de deslocar do estado i para o j , $P(i \rightarrow j)$ (Note que isto implica que $\sum_j P(i, j) = \sum_j P(i \rightarrow j) = 1$, i.e., a soma das linhas da matriz é igual a 1).

A equação de Chapman – Kolmogorov torna-se $\pi(t+1) = \pi(t)P$.

Usando a forma matricial, nós imediatamente vemos como a equação de Chapman – Kolmogorov funciona:

$$\pi(t) = \pi(t-1)P = (\pi(t-2)P)P = \pi(t-2)P^2$$

Equação 3-7

Continuando deste modo, mostra-se que:

$$\pi(t) = \pi(0)P^t$$

Equação 3-8

Definindo a n -ésima probabilidade de transição $p_{i,j}^n$ como a probabilidade de o processo está no estado j dado que ele iniciou no estado i , n instantes atrás, i.e.,

$$p_{i,j}^n = \Pr(X_{t+n} = s_j \mid X_t = s_i)$$

Equação 3-9

, segue imediatamente que $p_{i,j}^n$ é o ij -ésimo elemento de P^n .

3.3 METODOLOGIA DE SIMULAÇÃO MCMC

Levou aproximadamente 40 anos para MCMC penetrar no fluxo principal da prática estatística. Originado na literatura física estatística, foi bastante utilizado por algumas décadas em estatística espacial e análise de imagens. Nos últimos anos, MCMC tem tido um profundo efeito na estatística Bayesiana e tem também tido aplicações na estatística clássica. Recentes pesquisas têm aumentado consideravelmente os campos de aplicação dessa metodologia.

Nas últimas décadas, os estatísticos têm aumentado o uso da metodologia MCMC para simular distribuições complexas e não paramétricas de probabilidade. De acordo com Gilks et al (1998), o algoritmo de Gibbs é um dos mais conhecidos destes métodos e seu impacto na estatística Bayesiana, seguindo o trabalho de Tanner e Wong (1987) e Gelfand e Smith (1990), tem sido imenso como detalhado em muitos artigos, por exemplo, Smith e Roberts (1993), Tanner (1993), e Chib e Greenberg (1993). Uma considerável quantia de atenção está sendo dedicada agora ao algoritmo *Metropolis – Hastings (M-H)*, o qual foi desenvolvido por Metropolis, Rosenbluth, Rosenbluth, Teller, e Teller (1953) e subseqüentemente generalizado por Hastings (1970). Este algoritmo é extremamente versátil e tem como um dos casos especiais mais importantes o *amostrador de Gibbs*. O algoritmo M-H tem sido extensivamente usado no ramo da física, ainda embora que o artigo de Hastings, tenha sido pouco conhecido pelos estatísticos até recentemente.

Pelo fato de sua grande utilidade, aplicações do algoritmo M-H estão aparecendo continuamente na corrente literatura. Embora sua importância seja óbvia, não está disponível na literatura uma exposição simples ou intuitiva do algoritmo M-H, quando comparado a exposição de Casella e George (1992) para o amostrador de Gibbs.

A metodologia usual para a teoria markoviana em um espaço amostral de variáveis aleatórias contínuas inicia-se com um kernel de transição $P(x, A)$ para todo $x \in R^d$. O **kernel de transição** é uma função de distribuição condicional que representa a probabilidade de mover de x para um ponto qualquer no conjunto A .

Uma das maiores preocupações da teoria Markoviana é determinar condições sob as quais exista uma distribuição invariante, também conhecida como **distribuição meta**, que aqui designaremos por π^* e sob as quais iterações do **kernel de transição** convirja para a distribuição invariante. A distribuição invariante satisfaz:

$$\pi^*(dy) = \int_{R^d} P(x, dy)\pi(x)dx$$

Equação 3-10

, onde π é a densidade com respeito à medida de Lebesgue de π^* . A n -ésima iteração é dada por $P^{(n)}(x, A) = \int_{R^d} P^{(n-1)}(x, dy)P(y, A)$, onde $P^{(1)}(x, dy) = P(x, dy)$.

Na simulação MCMC, a distribuição meta $\pi(\cdot)$ é conhecida. É dessa densidade meta que amostras são *desejadas*. Entretanto, o **kernel de transição** não é conhecido. Para gerar amostras de $\pi(\cdot)$, os métodos encontram e utilizam um kernel de transição $P(x, dy)$, do qual a n

- ésima iteração converge para $\pi(\cdot)$ quando $n \rightarrow \infty$. O processo é iniciado em um arbitrário x e iterado um grande número de vezes. Depois deste largo número, as observações geradas a partir da simulação seguem aproximadamente a distribuição meta.

O problema, portanto, é encontrar um apropriado *kernel de transição* $P(x, dy)$, o que pode parecer ser uma procura por uma proverbial agulha no palheiro, o que é um tanto simplificado pelas seguintes considerações. Suponha que o kernel de transição, para alguma função $p(x, y)$, é expresso como:

$$P(x, dy) = p(x, y)dy + r(x)\delta_x(dy)$$

Equação 3-11

, onde $p(x, x) = 0$, $\delta_x(dy) = 1$ se $x \in dy$ e 0 de outra maneira, e $r(x) = I - \int_{R^d} p(x, dy)dy$ é a probabilidade que a cadeia permaneça em x . Da possibilidade que $r(x) \neq 0$, fica claro que a integral de $p(x, dy)$ ao longo de y não é necessariamente 1.

Agora, se a função $p(x, dy)$ na **Equação 3-10** satisfaz a condição de reversibilidade:

$$\pi(x)p(x, y) = \pi(y)p(y, x)$$

Equação 3-12

, então $\pi(\cdot)$ é a densidade invariante de $P(x, \cdot)$.

Intuitivamente, o lado esquerdo da condição de reversibilidade é a probabilidade não - condicional de mover de x para y , onde x é gerado de $\pi(\cdot)$, e o lado direito é a probabilidade não - condicional de mover de y para x , onde y também é gerado de $\pi(\cdot)$. A condição de reversibilidade diz que os dois lados são iguais.

O resultado nos dá uma condição suficiente (reversibilidade) que deve ser satisfeita por $p(x, y)$. Nós, mais adiante, iremos mostrar como o algoritmo Metropolis – Hastings encontra uma função $p(x, y)$, com esta propriedade, mas antes iremos expor o Método de geração de variáveis aleatórias conhecido como Método de Aceitação-Rejeição.

3.4 O MÉTODO DE ACEITAÇÃO – REJEIÇÃO (A-R)

O objetivo de despender um pouco de tempo com o método (A-R) é que o mesmo possui algumas semelhanças com o algoritmo Metropolis – Hastings (M-H), o qual é o nosso objeto de estudo. Entende-se aqui que o relato de tal método será bastante válido para total compreensão do algoritmo (M-H).

Suponha, então, que nós temos um método qualquer disponível para amostrar uma variável aleatória tendo uma função de densidade de probabilidade $g(x)$, este método pode ser um dos disponíveis na literatura, tal como o método da inversão, composição, convolução ou o método da transformação, para maiores detalhes veja Ross (2002). Nós podemos usar tal método como base para geração de uma variável aleatória contínua tendo função de densidade $f(x)$, gerando Y de g e então aceitando este valor gerado (Y) com uma probabilidade proporcional a $f(Y) / g(Y)$.

Especificamente, permita c ser uma constante tal que:

$$\frac{f(y)}{g(y)} \leq c, \text{ para todo } y$$

Nós, então, temos a seguinte técnica ilustrada na Figura 3-1 para gerar uma variável aleatória tendo densidade f .

O método A – R pode ser representado de forma algorítmica da seguinte maneira:

- Passo 1: Gere Y tendo densidade g .
- Passo 2: Gere um número aleatório U .
- Passo 3: Se $U \leq \frac{f(Y)}{cg(Y)}$, faz-se $X = Y$. Senão, retorne ao Passo 1.

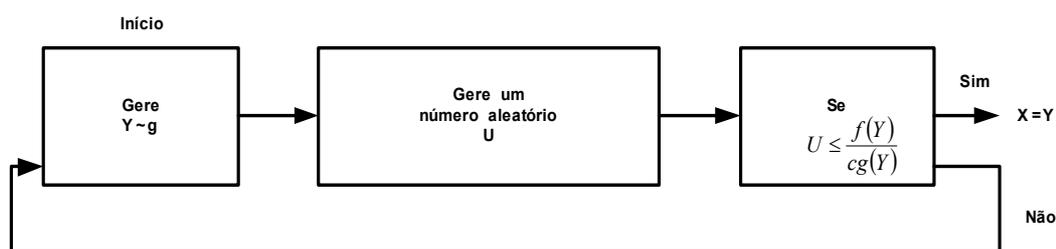


Figura 3-1– Método Aceitação – Rejeição

O método descrito aqui diz respeito à geração de variáveis aleatórias contínuas, mas todas as considerações feitas também valem para o caso discreto, somente diferindo que as funções de densidade de probabilidade são substituídas por distribuições de probabilidade no caso discreto.

A similaridade entre o método M-H e o A-R está no fato que ambos amostram variáveis aleatórias candidatas que podem ser aceitas ou rejeitadas com determinada probabilidade, diferentemente das técnicas clássicas de simulação, como o método da inversão, para maiores detalhes veja Ross (2002), nos quais amostras da variável aleatória são deterministicamente aceitas, com probabilidade 1.

3.5 O ALGORITMO METROPOLIS – HASTINGS (M – H)

Como no *método de Aceitação – Rejeição*, descrito na seção anterior, suponha que nós temos uma densidade que pode gerar candidatas. Desde que estejamos trabalhando com cadeias de Markov, entretanto, nós permitimos que esta densidade dependa apenas do estado corrente do processo. Portanto, a *densidade candidata* é denotada por $q(x, y)$.

Esta densidade deve ser interpretada como dizendo que quando um processo está no ponto x , a densidade gera um valor y de $q(x, y)$. Caso aconteça que a própria $q(x, y)$ satisfaça a condição de reversibilidade, dada pela Equação 3-12 para todo x, y , nossa pesquisa por um kernel de transição está finalizada. Entretanto, o mais provável e comum é que isso não seja verdade. Nós podemos encontrar, por exemplo, que para algum x, y ,

$$\pi(x)q(x, y) > \pi(y)q(y, x)$$

Equação 3-13

Neste caso, expressando-se de forma coloquial, o processo desloca-se de x para y mais freqüentemente do que de y para x . Um conveniente caminho para corrigir essa condição é reduzir o número de deslocamentos de x para y introduzindo uma probabilidade $\alpha(x, y) < 1$. Nós iremos nos referir a $\alpha(x, y)$ como a *probabilidade de deslocamento*. Caso o deslocamento não seja feito, o processo novamente retorna x como valor da distribuição meta, i.e., a cadeia de Markov não muda de estado. (Aqui está o contraste do algoritmo M-H com o método A-R no qual, quando um y é rejeitado, um novo par (y, u) é amostrado independentemente do valor anterior de y). Assim transições de x para y ($y \neq x$) são feitas de acordo com o seguinte:

$$p_{MH}(x, y) \equiv q(x, y)\alpha(x, y) \quad x \neq y$$

Equação 3-14

, onde $\alpha(x, y)$ ainda não foi determinado.

Considere novamente a inequação dada pela **Equação 3-13**. Ela informa-nos que o deslocamento de y para x não é feito da forma adequada. Portanto, nós definiríamos $\alpha(y, x)$ para ser tão largo quanto possível, logo, desde que se trata de uma probabilidade, seu limite superior é 1. Mas, agora a **probabilidade de deslocamento** $\alpha(x, y)$ é determinada requerendo que $p_{MH}(x, y)$ satisfaça a condição de reversibilidade, por isso então:

$$\pi(x)q(x, y)\alpha(x, y) = \pi(y)q(y, x)\alpha(y, x)$$

Equação 3-15

Sendo $\alpha(y, x) = 1$, temos que:

$$\pi(x)q(x, y)\alpha(x, y) = \pi(y)q(y, x)$$

Equação 3-16

Nós agora vemos que $\alpha(x, y) = \pi(y)q(y, x) / \pi(x)q(x, y)$. É óbvio que se a inequação em **Equação 3-13** for o contrário, nós fixamos $\alpha(x, y) = 1$ e derivamos $\alpha(y, x)$ como acima. As probabilidades $\alpha(x, y)$ e $\alpha(y, x)$ são, portanto, introduzidas para assegurar que os dois lados da **Equação 3-13** sejam balanceados ou, em outras palavras, que $p_{MH}(x, y)$ passe a satisfazer a condição de reversibilidade. Assim, nós mostramos que para que $p_{MH}(x, y)$ seja reversível, a **probabilidade de deslocamento** deve ser configurada, como segue:

$$\alpha(x, y) = \min \left[\frac{\pi(y)q(y, x)}{\pi(x)q(x, y)}, 1 \right]$$

Equação 3-17

Várias observações a respeito deste algoritmo merecem ser feitas. Primeiro, o algoritmo M-H está especificado por sua **densidade candidata** $q(x, y)$ da qual trataremos mais adiante. Segundo, se um valor candidato é rejeitado, o corrente valor é tomado como o próximo item na seqüência. Terceiro, o cálculo de $\alpha(x, y)$ não requer conhecimento da constante de normalização de $\pi(\cdot)$ porque ela aparece tanto no numerador quanto no denominador. Quarto, se a densidade candidata é simétrica, um importante caso especial é que $q(x, y) = q(y, x)$ e $\alpha(x, y)$ (**probabilidade de deslocamento**) fica reduzida a $\pi(y) / \pi(x)$; Logo, se a cadeia desloca-se para y , então, ela se desloca com probabilidade dada por $\pi(y) / \pi(x)$.

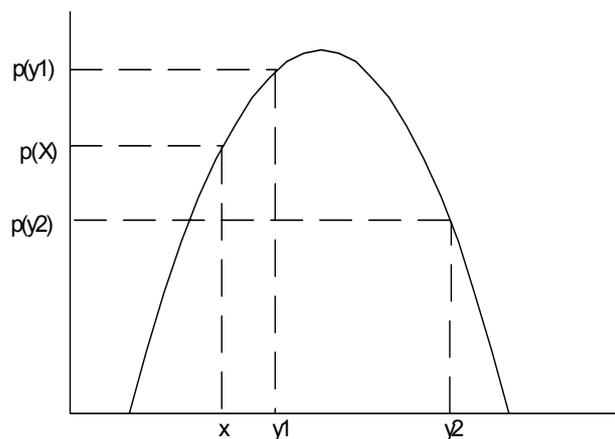


Figura 3-2 – Calculando Probabilidades de deslocamento com Função geradora de candidatas simétrica

Analisando a **Figura 3-2**, tem-se que do corrente ponto x , um deslocamento para o candidato y_1 é feito com certeza, enquanto um movimento para um candidato y_2 é feito com probabilidade $p(y_2)/p(x)$. Este é o algoritmo proposto por Metropolis et al. (1953). Interessantemente, também forma a base para vários algoritmos de otimização, notavelmente o método de ‘simulated annealing’.

Nós agora mostramos o algoritmo M-H em uma forma algorítmica iniciada com um valor arbitrário $x^{(0)}$:

- **Repita para $j = 1, 2, \dots, N$.**
- **Amostre y de $q(x^{(j)}, \cdot)$ e u de $U(0, 1)$.**
- **Se $u \leq \alpha(x^{(j)}, y)$**
 - **Faça $x^{(j+1)} = y$.**
- **Senão**
 - **Faça $x^{(j+1)} = x^{(j)}$.**
- **Retorne os valores $\{x^{(1)}, x^{(2)}, \dots, x^{(N)}\}$.**

Como em qualquer método MCMC, as amostras desenhadas são consideradas como uma amostra da densidade meta $\pi(x)$ somente após a cadeia ter passado pelo estágio transiente e o efeito do valor inicial fixado ter se tornado tão pequeno que pode ser ignorado. Na verdade, esta convergência para distribuição invariante ocorre sob condições de regularidade moderadas. As condições de regularidade requeridas são a irredutibilidade e a aperiodicidade dos dados amostrais. O que estas significam, respectivamente, é que, se x e y estão no

domínio de $\pi(\cdot)$, deve ser possível mover de x para dy em um finito número de iterações com probabilidade diferente de zero, e o número de deslocamentos requeridos para mover de x para y não requer que seja um múltiplo de algum inteiro.

Estas condições, entretanto, não determinam a taxa de convergência. Há, então, uma questão empírica em torno de qual o tamanho de uma amostra inicial n_0 que deveria ser descartada e durante quanto tempo a amostragem deveria ser simulada.

3.6 ESCOLHA DA DISTRIBUIÇÃO CANDIDATA $q(X, Y)$

Para implementar o algoritmo M-H, é necessário que uma *densidade candidata* satisfatória $q(x, y)$ seja especificada. Tipicamente, esta densidade é selecionada de uma família de distribuições que requer a especificação de parâmetros sintonizados, tais como o parâmetro de forma e de escala da distribuição. Trabalhos recentes têm dedicado considerável atenção a como estas escolhas devem ser feitas e embora a teoria esteja longe de completa, suficiente é conhecido para executar a maioria dos estudos práticos de simulação.

Uma família de densidades candidatas, que aparece no trabalho de Metropolis et al. (1953), é dada por $q(x, y) = q_1(y - x)$, onde $q_1(\cdot)$ é uma densidade multivariada. O valor candidato y é assim amostrado conforme o processo $y = x + z$, onde z é denominada variável aleatória incremental e segue a distribuição q_1 . Pelo fato da candidata ser igual ao corrente valor mais um ‘distúrbio’, este caso é chamado de uma cadeia *random walk*. Possíveis escolhas para q_1 incluem a densidade normal multivariada e a multivariada – t com os parâmetros especificados conforme os princípios descritos abaixo. Note que quando q_1 é simétrica, $q(x, y) = q(y, x)$, circunstância usual, a probabilidade de deslocamento então se reduz a:

$$\alpha(x, y) = \min \left\{ \frac{\pi(y)}{\pi(x)}, 1 \right\}$$

Equação 3-18

Uma segunda família de densidades candidatas é dada pela forma $q(x, y) = q_2(y)$ [veja Hastings (1970)]. Em contraste com o *random walk*, as candidatas são amostradas independentemente da corrente posição x , trata-se da chamada *cadeia independente*, utilizando a terminologia introduzida por Tierney(1994). Como no caso anterior, permite-se,

também, que q_2 seja uma densidade normal multivariada ou multivariada-t, mas agora é necessário especificar a posição da densidade candidata assim como sua dispersão.

Uma terceira opção, a qual parece ser uma solução eficiente quando disponível, é explorar a forma conhecida de $\pi(\cdot)$ para especificar uma densidade candidata. Por exemplo, se $\pi(t)$ pode ser escrito como $\pi(t) \propto \psi(t)h(t)$, onde $h(t)$ é uma densidade que pode ser amostrada e $\psi(t)$ é uma uniforme truncada, então configura-se $q(x, y) = h(y)$ (como em uma cadeia independente) para amostrar candidatas. Neste caso, a probabilidade de deslocamento requer somente o cálculo da função ψ (nem π nem h) e é dado por:

$$\alpha(x, y) = \min \left\{ \frac{\psi(y)}{\psi(x)}, 1 \right\}$$

Equação 3-19

Uma quarta alternativa para a densidade candidata $q(x, y)$ é a denominada **random walk em escala logarítmica**. Este tipo de densidade candidata é utilizado para componentes unidimensionais positivos, tais como, taxas de falha ou tempos de falha. A densidade $q(x, y)$ é escolhida tal que $\ln(y) - \ln(x)$ tenha uma densidade que é simétrica em torno de zero. Ou seja,

$q(x, y)$ tenha a seguinte forma: $q(x, y) = \frac{1}{\sqrt{2\pi\sigma^2}y} e^{-\frac{1}{2\sigma^2}(\ln(y/x))^2}$. Desse modo, a razão

correspondente a probabilidade de aceitação torna-se: $\frac{\pi(y)y}{\pi(x)x}$.

Nós agora retornamos para a crítica questão de escolher a dispersão ou escala da densidade candidata. Este é um importante assunto que tem aplicações para eficiência do algoritmo. A dispersão da densidade candidata afeta o comportamento da cadeia em no mínimo duas dimensões: uma é a “taxa de aceitação” (a porcentagem de vezes que um movimento para um novo ponto é feito), e o outro é a região do espaço amostral que é coberto pela cadeia. Para ver o porquê, considere a situação na qual a cadeia convergiu e a densidade está sendo amostrada ao redor da moda. Portanto, se a dispersão é extremamente larga, algumas das candidatas geradas estarão longe do corrente valor e, portanto, terão uma baixa probabilidade de serem aceitas. Reduzindo a dispersão corrigirá este problema, mas se a dispersão é escolhida muito pequena, tomará valores mais difíceis para atravessar o suporte da densidade, e regiões de baixa probabilidade serão subamostradas.

Recentes trabalhos realizados por Gilks et al (1998) discutiram este assunto no contexto de q_1 (a densidade proposta **random walk**). Eles mostraram que se as densidades proposta e meta

são normais, então a escala da última deveria ser sintonizada tal que a taxa de aceitação seja aproximadamente 0.45 em problemas unidimensionais e aproximadamente 0.23 em problemas cujo o número de dimensões é aproximadamente infinito, em quais a taxa de aceitação ótima gira em torno de 0.25 em problemas de seis dimensões, por exemplo. Isto é similar às recomendações de Muller (1993) apud Gilks et al (1998), que discute que a taxa de aceitação deveria girar em torno de 0.5 para a *cadeia random walk*.

4 ESTUDO DE CASO

4.1 Elicitação da opinião do Especialista

Um das maiores vantagens do Teorema de Bayes é permitir que o know-how implícito de um especialista seja utilizado em atualizações sucessivas do conhecimento sobre determinado assunto. Um especialista pode ser definido como uma pessoa idônea que possui muito treinamento e conhecimento em algum especial campo (Ayyub, 2001). O especialista é o fornecedor de opinião no processo de elicitação da opinião do especialista. Uma opinião de especialista pode ser definida como o julgamento formal de um especialista no qual o julgamento dele ou dela é buscado. Também, o julgamento pode significar uma crença que é baseada na incerteza da informação (Ayyub, 2001).

Serão apresentados a seguir alguns métodos que buscam extrair a opinião do especialista da maneira mais eficiente possível de modo que esta seja a mais fidedigna e próxima da realidade.

4.1.1 Métodos de Elicitação da Opinião do Especialista

Esta seção fornece uma síntese de alguns métodos que podem ser usados para elicitação de opiniões de especialistas dentro do contexto de Eng^a de Confiabilidade. Para aumentar as chances de sucesso no uso de métodos de elicitação, Cooke (1991) sugeriu algumas práticas relevantes, quais sejam:

1. Os assuntos ou questões devem ser claramente declarados para evitar qualquer ambigüidade. Algumas vezes pode ser necessário um teste dos temas ou questões para assegurar a adequada interpretação pelos indivíduos elicitados;
2. As questões ou assuntos devem ser determinados usando formato apropriado com respostas listadas, talvez graficamente expressas, para facilitar e otimizar o processo de elicitação;
3. Os analistas devem estar presentes durante a elicitação;
4. Treinamento e calibração dos especialistas devem ser executados. Exemplos devem ser apresentados com explanações da elicitação e conseqüentes agregação e consecução de resultados. Os analistas devem evitar instruir os especialistas ou conduzi-los a certas visões e respostas;

5. As seções de elicitação não devem ser muito alongadas. Para tratar muitos assuntos, várias seções com apropriados intervalos podem ser necessárias.

Elicitação indireta

O método de elicitação indireta é popular entre os teóricos da área. O método indireto é baseado na determinação de taxas pelos especialistas para alcançar um ponto de indiferença entre as opiniões apresentadas relacionadas a um determinado tema.

Outras técnicas de elicitação indireta foram desenvolvidas por pesquisadores para elicitar probabilidades a especialistas não familiarizados com questões estatísticas e de probabilidade. Por exemplo, analistas usaram a estimativa do tempo até falhar ou a idade de substituição de determinado equipamento como uma estimativa indireta da probabilidade de falha.

Exemplo: *Determinando taxas para propósitos de elicitação*

A determinação de taxas pode ser usada para avaliar subjetiva e indiretamente a probabilidade de ocorrência de um evento A , denominada $p(A)$. De acordo com este método, para um especialista E é nomeado um ticket de loteria da seguinte forma:

O Especialista E recebe R\$100 se A ocorrer

O interesse daqui por diante torna-se o valor que o especialista anexa a este ticket de loteria. Para uma assumida quantia de dinheiro $R\$x$, menor que $R\$100$, o especialista é perguntado se quer negociar o ticket pela quantia $R\$x$. A quantia x é incrementada até que um ponto de indiferença seja estabelecido, i.e., o ticket de loteria tenha a mesma utilidade que a quantia oferecida x . A posição $R\$x$ é denominada *equivalente certo* para o ticket de loteria.

Assumindo que o especialista é um indivíduo racional e não viesado, a posição x a qual é o equivalente certo para o ticket de loteria, fornece uma avaliação de um valor esperado. A utilidade esperada do ticket de loteria pode ser expressa como:

$$\text{Utilidade esperada do ticket de loteria} = R\$100kp(A)$$

Equação 4-1

, onde $p(A)$ = probabilidade de ocorrência de A e k = uma constante que representa a utilidade para o dinheiro como julgado pelo especialista. A função utilidade para o dinheiro pode ser uma função não – linear da quantia associada. Na posição do equivalente certo, $R\$x$ tem uma utilidade $k\$x$, a qual é equivalente ao valor esperado da utilidade do ticket de loteria como mostrado na **Equação 4-1**. Portanto, a seguinte condição pode ser fixada:

$$(\$100)k(p(A)) = k(\$x)$$

Resolvendo para $p(A)$ fornece:

$$p(A) = \frac{\$x}{\$100}$$

A utilidade do dinheiro no exemplo acima foi assumida ser linear, enquanto que a evidência empírica sugere que é altamente não - linear. De acordo com Ayyub (2001), Galanter (1962) construiu a tabela a seguir perguntando sobre a seguinte questão:

“Suponha que nós demos a você x reais; reflita sobre quanto feliz você está. Quanto nós teríamos que dá a você para que você ficasse duas vezes mais feliz?”

Tabela 4-1 – Dinheiro requerido para duplicar a felicidade

Dinheiro requerido para duplicar a felicidade		
Dinheiro dado	Duas vezes mais feliz	
X (R\$)	Média	Mediana
10	53	45
20	538	350
1000	10220	5000

Fonte: Ayyub, 2001

A seguinte função utilidade U foi desenvolvida baseada nos dados:

$$U(x) = 3.71x^{0.43}$$

É evidente que a vontade de uma pessoa correr um risco não aumenta linearmente com o aumento de x . Essa é uma das principais desvantagens do método de elicitación indireta.

Testes similares foram executados para perdas de dinheiro e a relação da mesma com a infelicidade, porém foram inconclusivos.

Método Direto

Este método elicitava estimativas diretas do grau de crença de um especialista em determinado assunto. Embora seja de natureza simples, este método pode produzir resultados errôneos, especialmente de especialistas que não estão familiarizados com a noção de probabilidade. O método Delphi e a técnica de grupo nominal fazem parte dessa categoria.

Estimativa Paramétrica

A estimativa paramétrica é usada para avaliar intervalos de confiança em um parâmetro de interesse tal como um valor médio. O processo de estimação pode ser executado na forma de um procedimento composto por dois passos, como segue (Cooke, 1991):

- i) Obtenha uma estimativa mediana de uma probabilidade (m), e
- ii) Obtenha a probabilidade (r) que o valor verdadeiro excederá em 10 vezes o valor mediano (m).

Os valores m e r podem ser usados para computar as fronteiras de confiança para os percentis de 5% e 95%.

4.2 Função de Verossimilhança

Um primeiro passo em qualquer procedimento de inferência estatística, seja clássico ou Bayesiano envolve o desenvolvimento da função de Verossimilhança. Como já relatado a função de verossimilhança constitui a probabilidade de que a evidência E seja observada se algum parâmetro de interesse é verdadeiro (ocorre).

Para determinação de uma função de verossimilhança deve-se analisar a natureza dos dados disponíveis, os quais representam a Evidência, para notar se os mesmos adequam-se a uma densidade de probabilidade paramétrica.

Para o nosso caso de estudo, a evidência é representada por um conjunto de falhas no tempo oriundo de uma população homogênea de equipamentos genéricos que aqui iremos designar por W .

População homogênea de equipamentos W :

$$1) W_1 : \langle k_1, T_1 \rangle;$$

$$2) W_2 : \langle k_2, T_2 \rangle;$$

$$3) W_3 : \langle k_3, T_3 \rangle;$$

$$4) W_4 : \langle k_4, T_4 \rangle;$$

$$5) W_5 : \langle k_5, T_5 \rangle;$$

.....

$$n) W_n : \langle k_n, T_n \rangle, \text{ onde } k_i \text{ representa o número de falhas observadas,}$$

T_i o tempo total no qual se observaram k_i falhas e

n é o n° total de equipamentos.

Em nosso contexto de aplicação os dados que vão ser utilizados na análise são homogêneos, i.e., todos os equipamentos W_i tem o mesmo comportamento quanto a taxa de falha λ . Devido a isso, podemos agrupar os dados fornecendo o seguinte:

$$k = \sum_{i=1}^n k_i$$

Equação 4-2

$$T = \sum_{i=1}^n T_i$$

Equação 4-3

, o que caracteriza um processo de amostragem de Poisson, onde tem-se k falhas durante um período de tempo T , fornecendo uma *taxa de falha empírica*.

Uma variável aleatória X segue a distribuição de Poisson com parâmetro $\lambda > 0$ se sua função de probabilidade é dada por:

$$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

Equação 4-4

, com o parâmetro λ sendo usualmente referido como a taxa de ocorrência. A notação utilizada será $X \sim Po(\lambda)$.

Portanto, a função de verossimilhança, a qual representará os dados de campo de falha empíricos será modelada por uma distribuição de Poisson com parâmetro λ .

Por conveniência matemática, introduziu-se a transformação monotônica $\lambda = -Ln(u)$. Daí a distribuição de probabilidade de Poisson pode ser expressa em termos da variável u , já que $0 < \lambda < \infty$ e $0 < u < 1$.

Portanto, a função de verossimilhança centrada na variável u , tal que $0 < u < 1$ é dada por:

$$P(X = k) = \frac{u(-Ln(u))^k}{k!} \quad k = 0, 1, 2, \dots$$

Equação 4-5

4.3 Distribuição a priori

Anteriormente, foram relatados alguns métodos disponíveis na literatura que podem ser usados para extrair o grau de confiança ou medida do conhecimento de um indivíduo (especialista) sobre uma dada proposição ou evento. O que se busca, portanto, é uma medida que representa o grau de conhecimento do especialista. Este grau de conhecimento é subjetivo por essência. Subjetivo, pois é uma representação do estado mental do indivíduo e não do mundo externo. Uma opinião é uma avaliação subjetiva ou estimativa da qualidade ou quantidade de algo de interesse que parece verdadeiro, válido ou provável (Notas de aula Droguett, 2002).

Ao invés de definir uma distribuição a priori para a taxa de falha $0 < \lambda < \infty$, podemos equivalentemente definir uma priori para a variável $0 < u < 1$. Concentrando em u , uma distribuição a priori que é: (i) matematicamente tratável, (ii) é definida ao longo do intervalo $[0;1]$ e (iii) não impõe outras restrições em u , é a *distribuição de probabilidade Beta*, cuja função de distribuição de probabilidade é dada por:

$$B(u | a, b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} u^{a-1} (1-u)^{b-1}$$

Equação 4-6

, onde “ a ” e “ b ”, ambos não – negativos, são os dois parâmetros da distribuição. Esses parâmetros determinam o formato da distribuição Beta:

- 1) Simétrica quando $a = b$;
- 2) Inclinada à direita quando $a < b$;
- 3) Inclinada à esquerda quando $a > b$;
- 4) Formato “ U ” quando $(a - 1)$ e $(b - 1)$ são negativos;
- 5) Formato “ J ” quando $(a - 1)$ é negativo.

O conhecimento do especialista será utilizado para estimar os parâmetros “ a ” e “ b ” da distribuição Beta. Van Dorp e Mazzuchi (2000) mostraram que os parâmetros de uma distribuição Beta podem ser resolvidos sob duas restrições representadas por um inferior e superior percentis e desenvolveram um procedimento numérico para calcular esses parâmetros. O método em Van Dorp e Mazzuchi (2000) resolve para um único conjunto de parâmetros de uma distribuição Beta que satisfaz as duas restrições de percentis. É provado, portanto a existência e a unicidade da solução.

Portanto, o objetivo do processo de elicitação do especialista corresponde a educação de dois percentis. Por conveniência de interpretação por parte do especialista, escolheu-se dois percentis bastante utilizados em engenharia, o 5% e 95% percentis. Estes percentis correspondem a uma estimativa otimista e pessimista, respectivamente, o que pode ser ilustrado através da Figura 4-1.

A estimativa *otimista* (α) corresponde a uma estimativa da taxa de falha mínima que um equipamento W pode ter. É obtida supondo-se condições totalmente favoráveis durante a operação do equipamento. A estimativa *pessimista* (β) é uma estimativa da taxa de falha máxima que o equipamento W pode possuir. Ocorre apenas em condições totalmente adversas. A possibilidade de eventos drásticos e catastróficos não é considerada, a menos que eles estejam claramente associados ao projeto (Moreira, 2001).

As estimativas *otimista* (α) e *pessimista* (β) correspondem respectivamente aos percentis de 5% e 95%, respectivamente. Portanto, temos duas restrições de percentis, que correspondem aos dois pares ordenados ($\alpha, 5\%$) e ($\beta, 95\%$). Fazendo uso do procedimento

numérico desenvolvido por Van Dorp e Mazzuchi (2000) é possível estimar os parâmetros da “a” e “b” distribuição Beta.

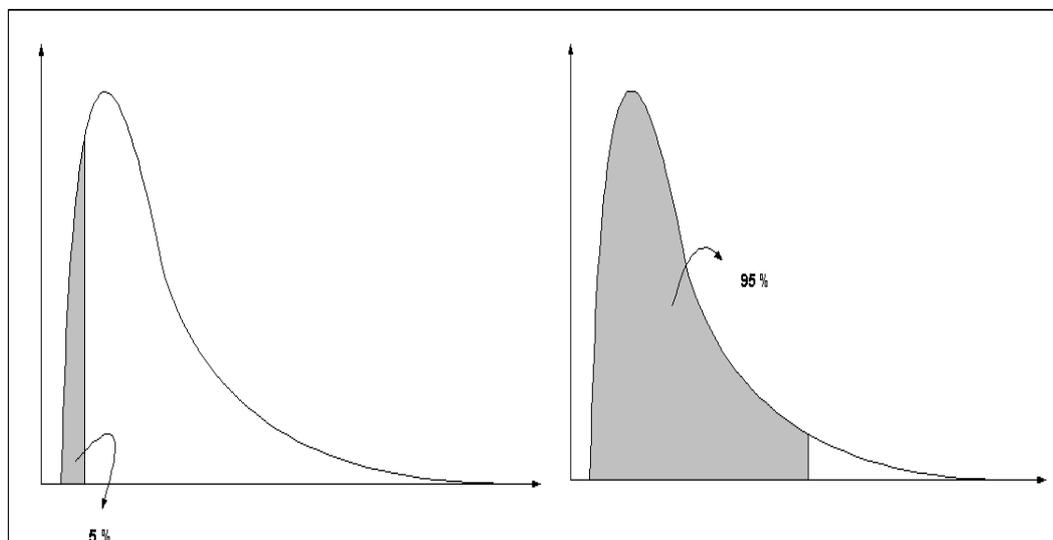


Figura 4-1 – Elicitando a opinião do especialista

4.4 Aproximação da distribuição a posteriori

A distribuição posteriori de u segue para função de verossimilhança dada por uma distribuição de probabilidade *Poisson* (Equação 4-5) e uma priori dada por uma distribuição *Beta* (Equação 4-6). Dado estas distribuições, a avaliação analítica da distribuição a posteriori não é possível. Daí, a necessidade para o uso da bem conhecida metodologia MCMC para posterior análise, mais especificamente do algoritmo Metropolis – Hastings.

O objetivo do método MCMC aqui é produzir amostras de u que tem como distribuição a distribuição meta $\pi(\cdot)$, que no nosso caso é dada por uma *BetaPoisson*. A alternativa escolhida como densidade candidata foi a “*random walk em escala logarítmica*”, devido ao fato de nosso parâmetro de interesse ser um componente positivo e unidimensional: taxa de falha.

Após as amostras de u serem geradas, determinam-se amostras para taxa de falha λ , que é nosso parâmetro de interesse, pela relação $\lambda = -Ln(u)$. Obtendo uma das métricas características de confiabilidade, por exemplo, a distribuição de probabilidade da taxa de falha (PDF), outras são obtidas por relação direta, tais como a Função de Distribuição Acumulada (CDF), entre outras.

5 RESULTADOS E DISCUSSÃO

O objetivo do nosso Estudo de Caso é a avaliação de métricas de confiabilidade a posteriori para a taxa de falha λ . É necessário para alcance de tal objetivo, portanto, elicitare a opinião do especialista em relação aos percentis de probabilidade de 5% e 95%, que correspondem às estimativas otimista e pessimista, respectivamente, para λ e de um conjunto de dados de falha do equipamento que é o nosso objeto de estudo.

No nosso caso esse equipamento corresponde a *Bombas de resfriamento utilizadas em uma Usina Nuclear*. Trata-se de diferentes bombas atuando em diferentes plantas de produção nuclear, mas que possuem o mesmo comportamento em relação à λ , e portanto podem ser tratados como equipamentos oriundos de uma mesma população, i.e., de uma população homogênea.

Os dados que representam os dados de campo, i.e., a evidência para tal estudo e a elicitare da opinião do especialista podem ser visualizados na Tabela 5-1 e Tabela 5-2, respectivamente:

Tabela 5-1 – Dados de falha empíricos

Planta	Nome do Equipamento	Nº de Falhas	Horas Operacionais
Zion	Centrifugal Charging Pump	0	7.60E+04
Zion	Service Water Pump	0	1.52E+05
IP-2	Component Cooling Pump	0	7.40E+04
IP-2	Service Water Pump	2	7.40E+04
IP-3	Component Cooling Pump	0	4.80E+04
IP-3	Service Water Pump	3	7.60E+04
BF	RHRW/EECW Pump	9	1.02E+05
Oconee	Service Water Pump	2	4.80E+04
	Total	16	6.50E+05

Fonte: Dados de teste

Como os dados de campo são homogêneos, podemos agrupá-los fornecendo um valor empírico para a taxa de falha λ , qual seja:

$$\lambda = \frac{16}{6.50E+05} = 2.46E-05 \text{ falhas / h}$$

Equação 5-1

Tabela 5-2 – Elicitação da Opinião do Especialista

Elicitação da Opinião do Especialista	
Percentil - 5% (Otimista)	Percentil - 95% (Pessimista)
2.03E-06 falhas/h	9.83E-05 falhas/h

Fonte: Especialista

Os dados da Tabela 5-2 foram obtidos através do método de elicitación direta, dado que o especialista elicitado tinha vastos conhecimentos estatísticos.

Com os dados obtidos na Tabela 5-2 e utilizando o procedimento desenvolvido por Van Dorp e Mazzuchi (2000), pode-se determinar quais são os parâmetros 'a' e 'b' da distribuição Beta, os quais estão na Tabela 5-3:

Tabela 5-3 – Parâmetros da Distribuição Beta

Parâmetros da Distribuição Beta	
a	9.54e-007
b	0.000975

Fonte: Autor

Determinados tais parâmetros e fazendo uso da **Equação 4-5**, é possível utilizar a metodologia MCMC, mais precisamente o algoritmo Metropolis – Hastings (M-H), para avaliação da distribuição a posteriori para a variável u , tal que $u = \exp(-\lambda)$ e $0 < u < 1$. Concentrando os esforços em u , por motivos de conveniência matemática e computacional, é possível após 'rodar' o algoritmo M-H, utilizar a transformação inversa $\lambda = -\ln(u)$ e assim determinar a distribuição a posteriori para taxa de falha λ .

Usando o algoritmo M-H, uma seqüência de comprimento $M = 100.000$ foi gerada para o Estudo de Caso descrito, fornecendo como *output*, a seqüência de variáveis $\underline{\lambda}_i^{seq} = \{\lambda_i^1, \dots, \lambda_i^{100000}\}$.

Destas $M = 100.000$ amostras nem todas são utilizadas para inferência da distribuição a posteriori para a taxa de falha λ . Isto é realizado com o objetivo de diminuir os efeitos do valor inicial da cadeia na análise final dos resultados. Do comprimento total M da cadeia, é retirada uma seqüência de comprimento L . Esta técnica é chamada de *Burn-In*. Existe pouca literatura sobre tal assunto, mas o suficiente para saber que é necessária. Atualmente, a determinação do comprimento L restringe-se na maioria dos casos à análise visual. No nosso caso, foram descartadas 10% de M , i.e., $L = 10.000$.

Para executar tais cálculos, devido à complexidade dos mesmos, foi desenvolvido um software em código fonte em linguagem C++ e interface em linguagem Java. A saída desse software pode ser ilustrada nas figuras abaixo, quais sejam:

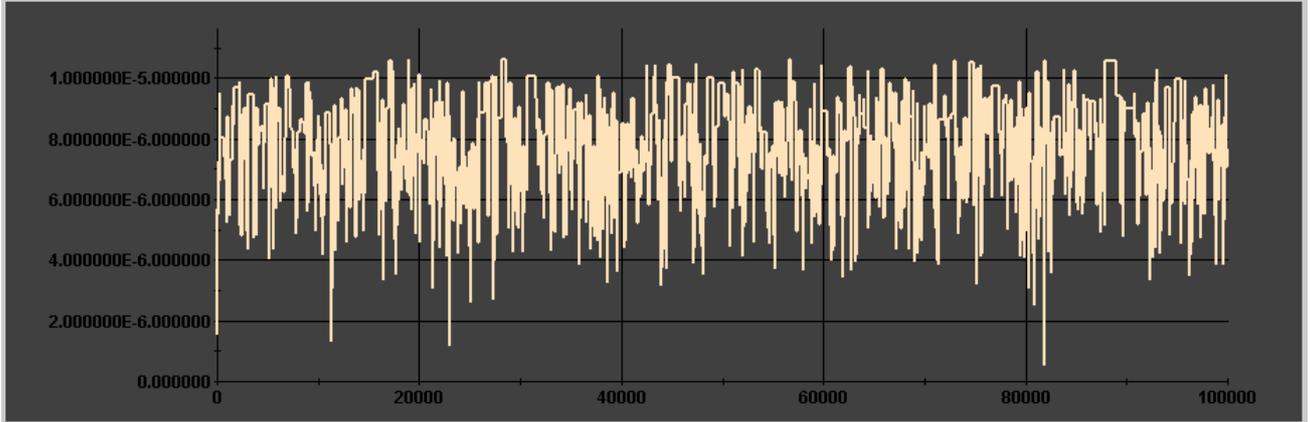


Figura 5-1 – Valores gerados para λ pelo algoritmo M-H

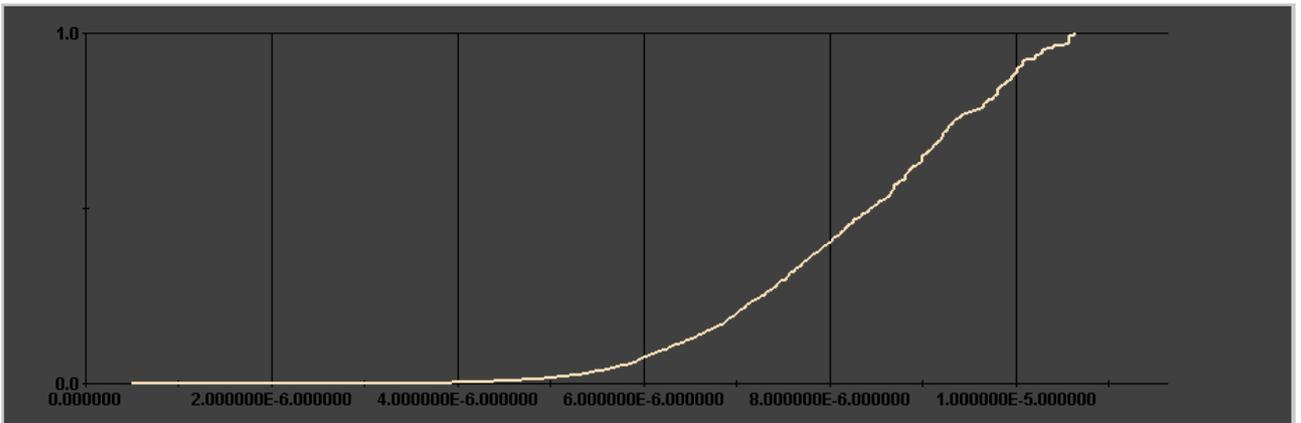


Figura 5-2 – Função de Distribuição Acumulada para λ (CDF)

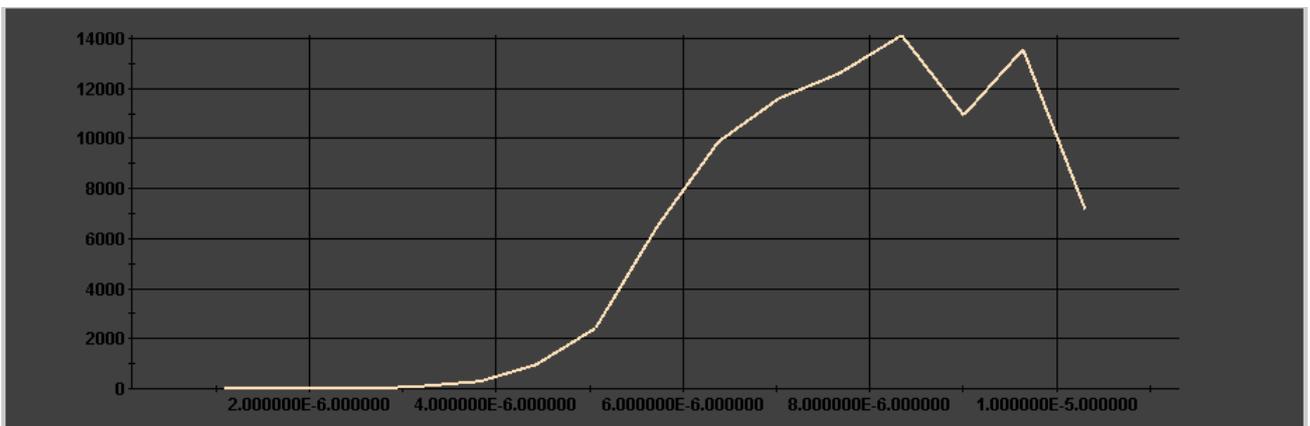


Figura 5-3 – Função de Densidade de Probabilidade para λ (PDF)

Obteve-se, portanto, a distribuição a posteriori não – paramétrica para a taxa de falha λ , representada pela Figura 5-3, para a Bomba de resfriamento da Usina Nuclear em questão. Como trata-se de uma PDF não paramétrica, tem – se valores médios de intervalos associados a frequências de ocorrência, onde nota-se através da Figura 5-3, que o valor mais freqüente ou valor modal de λ está em torno de $8.25E-06$ falhas/h. A média amostral simulada foi de $7.96E-06$ falhas/h, o que significa que o tempo médio para que *1 falha* ocorra é de aproximadamente *125698 horas operacionais*. O desvio padrão amostral, métrica que avalia a dispersão entre os dados, foi avaliado em $1.53E-06$ falhas / hora.

6 CONCLUSÕES

Este trabalho apresentou a metodologia de inferência Bayesiana para avaliação de métricas de confiabilidade. Esta metodologia surge como alternativa para a avaliação da confiabilidade de um equipamento, produto ou sistema de produção quando a escassez de dados se torna um obstáculo para tal. A grande motivação para o presente trabalho se dá pelas limitações das técnicas tradicionais que se baseiam no prisma frequentista para avaliação da Confiabilidade.

Em sistemas de produção do mundo real, onde é bastante comum que o tempo decorrido entre a concepção do produto e o seu lançamento no mercado seja demasiadamente curto, é praticamente impossível que grandes amostras de falha sejam geradas. Após o *ambiente da qualidade on-line* da organização ter determinado qual a tendência atual de mercado, o tempo até o produto chegar às mãos do consumidor não deve ser elevado. Portanto, a avaliação da confiabilidade exige que outras informações de falha sejam utilizadas, além dos dados empíricos.

Outro campo de aplicação bastante propício à inferência Bayesiana é a produtos altamente confiáveis, como as bombas de resfriamento do nosso Estudo de Caso, para os quais tempos médios até falhar (MTTFs) que excedem anos não são incomuns. Portanto, se a avaliação da confiabilidade para esses equipamentos fosse realizada através de uma metodologia não Bayesiana, grandes amostras de falha seriam necessárias o que levaria muito tempo e custo para serem adquiridas, tornando a análise inviável tanto técnica quanto economicamente.

A adaptação de outras fontes de informação de falhas como o conhecimento de especialistas surge como a grande vantagem do Teorema de Bayes. A capacidade de aglutinar diferentes tipos de dados de falha para obtenção da distribuição a posteriori para o parâmetro de interesse está permitindo que os conceitos da Teoria Bayesiana sejam utilizados em vários ramos da ciência.

É óbvia a grande vantagem da utilização de várias fontes de informação de diferentes naturezas para determinação da distribuição a posteriori para algum parâmetro de interesse, no nosso caso a taxa de falha λ . Porém em problemas do mundo real é praticamente impossível que a distribuição a posteriori possa ser avaliada de forma analítica. Daí entra a utilidade da metodologia MCMC, a qual também foi apresentada nesse trabalho, em problemas de inferência Bayesiana. A metodologia MCMC proporciona a geração de números aleatórios

utilizando conceitos de integração Monte Carlo e Cadeias de Markov. Dentre os algoritmos mais conhecidos da metodologia MCMC estão o Metropolis – Hastings (M-H), o qual foi utilizado no presente trabalho e o amostrador de Gibbs, que é um caso particular do algoritmo M-H.

No capítulo 1, foi apresentada a justificativa desse trabalho e os objetivos gerais e específicos do desenvolvimento do presente trabalho.

No capítulo 2, foi apresentada a metodologia que foi empregada mostrando o problema da inferência Bayesiana e fazendo um comparativo entre modelos clássico e bayesiano para avaliação de métricas de confiabilidade.

No capítulo 3, foi apresentada a utilidade da metodologia MCMC dentro do contexto do problema de inferência bayesiana, enfatizando o método de simulação Metropolis – Hastings.

No capítulo 4, foi apresentado o estudo de caso que corresponde a avaliação da taxa de falha de bombas que operam em um sistema de resfriamento de uma Usina Nuclear.

No capítulo 5, foram apresentados os resultados e discussão sobre tal estudo de caso.

A contribuição do presente trabalho para a minha formação enquanto Eng^o de Produção foi a oportunidade de congregiar temas relativos à linha de Pesquisa Análise de Risco e Eng^a de Confiabilidade, a qual participei de projetos de iniciação científica enquanto aluno de graduação e na qual pretendo realizar o mestrado acadêmico em Eng^a de Produção.

6.1 Desafios Futuros

Como sugestões de continuação do presente trabalho temos como alternativas:

- Avaliar a sensibilidade da opinião do especialista na determinação dos parâmetros da distribuição Beta;
- Extrapolar o método desenvolvido por Van Dorp e Mazzuchi (2000) para a distribuição LogNormal, que geralmente funciona muito bem na modelagem da distribuição a priori;
- Estender o software desenvolvido para Evidências de falha de outras naturezas, o que permitiria a avaliação de outras distribuições a posteriori de pares não conjugados;
- Adaptação da função de distribuição de probabilidade e/ou da distribuição acumulada através de curvas de Bézier;
- Realizar um Teste de Aderência com o objetivo de se determinar qual distribuição paramétrica de probabilidade melhor se adequa aos dados obtidos a partir de simulação.

REFERÊNCIAS BIBLIOGRÁFICAS

- 1) AYYUB, B. M. (2001)- Elicitation of Expert Opinions for Uncertainty and Risks. CRC Press. Florida.
- 2) COOKE, R. M. (1991)- *Experts in Uncertainty*. Oxford University Press, New York.
- 3) DEITEL, H. M. & DEITEL, P. J. (2001)- *Java™, como programar*. 3 ed. Bookman. Porto Alegre.
- 4) FIRMINO, P. R. (2004) – Redes Bayesianas para a Parametrização da Confiabilidade em Sistemas Complexos. Engenharia de Produção, Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências.
- 5) GEYER, C. J. (1992)- Practical Markov chain Monte Carlo. *Statistic Science*. Vol. 7, 473-511.
- 6) GILKS, W. R.; RICHARDSON, S. & SPIEGELHALTER, D. J. (1998)- *Markov chain Monte Carlo in Practice*. Chapman & Hall/CRC. Florida.
- 7) HASTINGS, W. K. (1970) – *Monte Carlo Sampling methods using Markov Chain and their applications*. *Biometrika*, 57, 97 – 109.
- 8) HOLZNER, S. (2002) - *C++ Black Book, Guia Completo para Resolução de Problemas*. São Paulo: Makron Books.
- 9) MAGALHÃES, M. N. & Lima, A. C. P. (2001)- *Noções de Probabilidade e Estatística*. 3 ed. Departamento de Estatística IME - USP. São Paulo.
- 10) MARTZ, H. F. & WALLER, R. A. (1982) - *Bayesian reliability analysis*. 2 ed. Krieger Publishing Company. Florida.
- 11) MODARRES, M.; KAMINSKIY, M., KRIVTSOV, V. (1999) - *Reliability Engineering and Risk Analysis*. Ed. Marcel Dekker. New York.
- 12) MOREIRA, D. A. (2001)- *Administração da Produção e Operações*. Ed. Pioneira. São Paulo – SP.

- 13) Notas de aula de DROGUETT, E., 2002 – Disciplina Eng^a de Confiabilidade, UFPE, Recife, não publicado.
- 14) PALADINI, E. P. (2000) – *Gestão da Qualidade*. Atlas. São Paulo, SP.
- 15) PRESS, W. H., FLANNERY, B.P., TEUKOLSKY S.A. And VETTERING W.T. (1989). *Numerical recipes in C++*, Cambridge University Press, Cambridge.
- 16) ROSS, S. M. (2002)- *Simulation*. 3 ed. Academic Press. California.
- 17) TIERNEY, L. (1994) – *Markov chains for exploring posterior distributions* (with discussion). *Ann.Statisc.*, 22, 1701 – 1762.
- 18) VAN DORP, J.R., MAZZUCHI, T.A., 2003. *Parameter Specification of the Beta Distribution and its Dirichlet Extensions Utilizing Quantiles*. *Beta Distribution and its Applications*, 29(1), 1-37 (2003).
- 19) VAN DORP, J.R., MAZZUCHI, T.A., 2000. *Solving for the parameters of Beta distribution under two quantile constraints*. *J. Statist. Comput. Simulation* 67, 189-201.
- 20) SUN (2001)- *JavaTM Programming Language*. Sun Microsystems. Colorado.