

## Phase-Type Distribution and Moment Matching

### Advanced Topics in Systems Performance Evaluation

Professors: Paulo Maciel/Ricardo Massa

Group 1

Eric Borba erb@cin.ufpe.br Erico Guedes eacg@cin.ufpe.br Jonas Pontes jcnp@cin.ufpe.br



Universidade Federal de Pernambuco 2015

Cln.ufpe.



- Preliminaries
- Phase-type Definition
- Moment Matching Methods



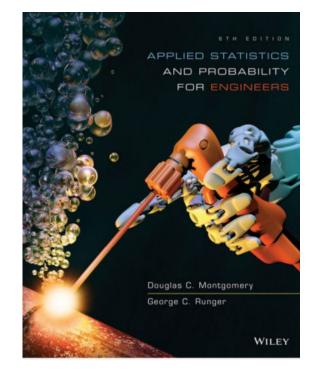
# "Statistics is a science that helps us make decisions and draw conclusions in the presence of variability".





## **Douglas Montgomery**





Cln.ufpe.



# "In practice, totally deterministic systems are unlikely due to influence of unpredictable factors."





## Theme:

## Phase-type Distribution And Moment Matching





## Theme:

Phase-type Distribution And Moment Matching

CIn.ufpe. 7







a scientific **test** in which you perform a series of actions and carefully observe their effects in order to learn about something [Merriam, 2015].

- **Experiment 1: toss a coin twice** 
  - Which are the possible outcomes? Ο

0.5 0.4

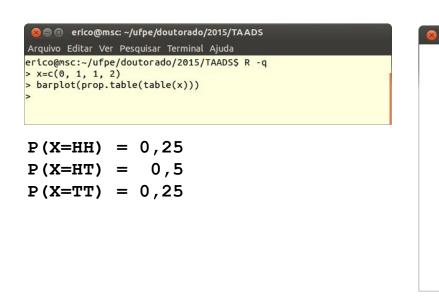
0.3

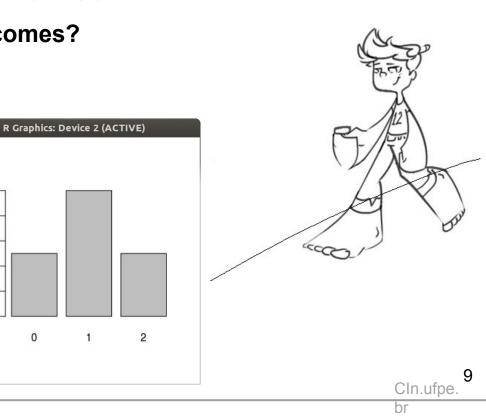
2 0

0.1 0.0

0

HH, HT, HT, TT Ο

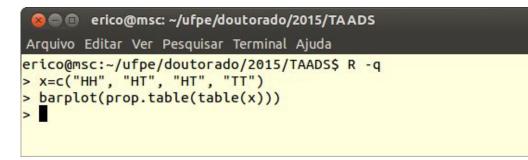


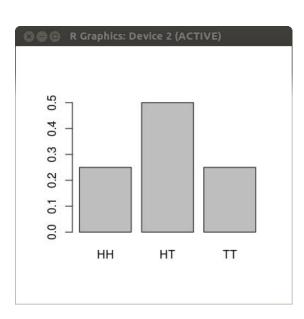




• Experiment 1: toss a coin twice

### • With Categorical Variable: HH, HT, HT, TT



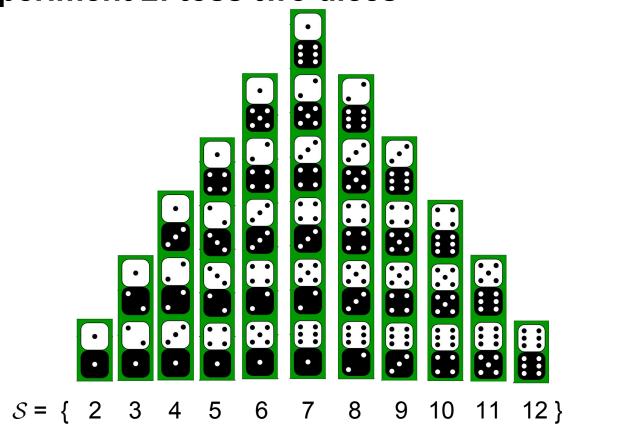


Cln.ufpe

10



Experiment 2: toss two dices



1 Cln.ufpe. br



• Exponential Distributions

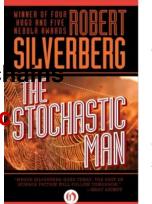


### It possesses the so-called memoryless property

it often leads to the

introduction of Markov c

in the study of stochastic systems.



"The concepts of cause and effect are fallacies. There are only seeming causes leading to apparent effects."

"A system in which events occur according to a law of probability but aren't individually determined in accordance with the principle of causality is a stochastic system."

Cln.ufpe

br



### • Stochastic: Merriam-Webster

Dictionary Thesaul	rus Medical Scrabble® Spanish Cent
stochastic	SEARCH
🚱 Games 🛛 🛗 Word of the Day 🛛 💽	Video 🍐 Blog: Words at Play 📑 My Faves
Games word of the Day	indeo biog. Words derindy
Dictionary	
	SAVE POPULARIT
Dictionary	SAVE POPULARIT
Dictionary stochastic 🖘	SAVE POPULARIT



### • Em bom Português...

MICHAELI	
Moderno Dicionário	Dicionário de Português Online
português	Significado de "estocástico"
Sobre o dicionário	dicionário
Gramática e curiosidades	Digite a palavra estocástico português buscar busca avançada
Gula prático da nova ortografia	lista por ordem alfabética: A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
Índice de verbetes	Compartilhe esta página: 🖾 🖪 Curtir 🕼 G+ Compartilhar Tweetar
inglês → português português → Inglês	estocástico es.to.cás.ti.co
Dicionário Escolar	adj (gr stokhastikós) 1 Relativo à estocástica. 2 Diz-se de processos que não
alemão → português Português → alemão	são submetidos senão a leis do acaso.



### • Exponential Distributions

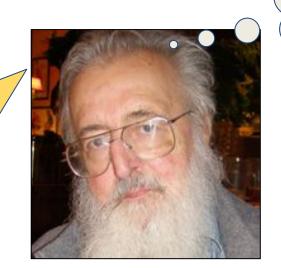
- The sole parameter  $\lambda$  of the exponential distribution can be interpreted explicitly under different circumstances.
- For instance:
  - λ is interpreted as the failure rate in reliability theory while it represents the arrival rate of the Poisson process



 Phase-type distributions were introduced by Neuts, in 1975, as a generalization of the exponential distribution.

Since then, matrix-analytic methods have become an **indispensable tool** in stochastic modeling and have found applications in the analysis and design of :

- manufacturing systems
- telecommunications networks
- risk/insurance models
- reliability models
- inventory and supply chain systems



**pioneered** in matrix-analytic methods in the study of queueing models



### ● PH-type Distributions → Motivations

- can be defined on Markov chains
- have a partial memoryless property
  - which often leads to Markovian models that are analytically and algorithmically tractable



• PH-type Distributions

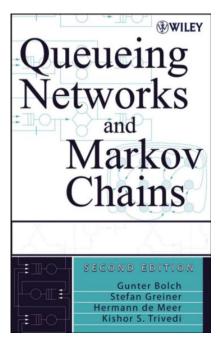
They constitute a versatile class of distributions that can approximate arbitrarily closely any probability distribution defined on the nonnegative real line



• Pointed references







Approximation problem:

 approximating general distributions by phasetype distributions



MASARYK UNIVERSITY FACULTY OF INFORMATICS



Phase-Type Approximation Techniques

BACHELOR THESIS

Zuzana Komárková

Brno, Spring 2012

19

Cln.ufpe.

br



• Keywords

Semi-Markov Process(SMP)

Continuous Time Markov Chain(CTMC)

**Complex Systems** 

**Poisson Process** 

Generalized Semi-Markov Process(GSMP)

**Stochastic Models** 









## • Continuous-Time Markov Chain(CTMC) Definition

• A CTMC is a tuple( $s, R, v_o$ ), where:

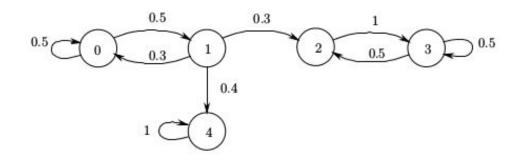
**S** is a finite set of states

- R: S x S →  $\mathbb{R}$  is a rate matrix, such that R(i,j) ≥ 0 for i ≠ j and R(i,j) = - $\sum_{k \neq i} R(i,k)$ , and
- $v_o$ : is an initial probability vector



### • CTMC - Absorbing state

- A state is called absorbing if once it has been reached it is impossible to leave it
- There is any absorbing state in below CTMC?





- Named from?
  - The name of phase-type distributions comes from the key idea behind it, which is to model waiting times as being made up of the time taken to move through a number exponentially distributed phases.
  - These phases are equivalent to states of the underlying Markov chain
    - The term phases are used when writing about PH-distributions, and states when writing about CTMC





### • Quantitative analysis of man-made systems like:

- computer systems
- communication networks
- manufacturing plants
- logistics networks
- **(...)**

# is often done by means of discrete event models that are analyzed numerically or by simulation

25



### • PHD - PHase-type Distribution

 The approximation of a general distribution by a PHD is a complex non-linear optimization problem for which only recently (2014)
 computational algorithms have been proposed which have not found their way into broadly available modeling software yet.





### • PHD - PHase-type Distribution

◦ Consider an absorbing CTMC with n+1 states, where n ≥ 1, such that states

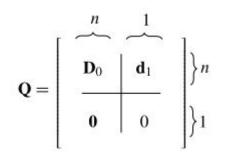
1, ..., n are transient and state 0 is an absorbing state.

- Further, let the chain have an initial probability of starting in any of the **n+1** phases given by the probability vector ( $p_0 p$ ).
- The (continuous) n-phase phase-type distribution (PH-distribution) is the distribution of time from the above process's starting until absorption in the absorbing state.



### • Continuous-Time Markov Chain(CTMC):

Infinitesimal Generator Matrix



 $\boldsymbol{D}_{0}$ : submatrix describing only transitions between transient states

 $\boldsymbol{d}_1$  : transitions from transient states to the absorbing state

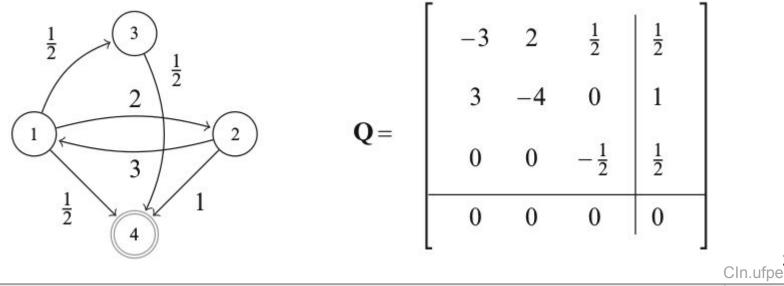
**0**: contains 0's(no transitions from the absorbing states to transitions states can occur)

28 Cln.ufpe. br



### • Continuous-Time Markov Chain(CTMC):

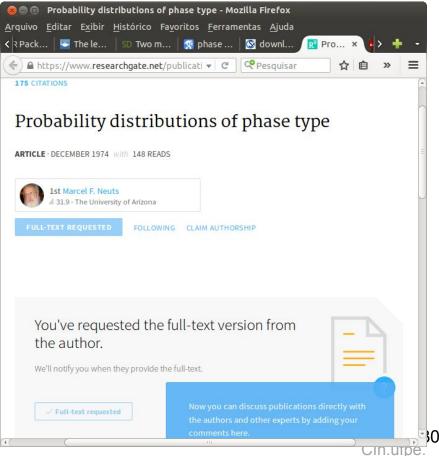
 $\circ$  Example: What is the Infinitesimal Generator Matrix **Q** of below CTMC?



29

## Definition - PH Distribution

• Neuts' definition





### • Neuts' definition

1981 Winter Simulation Conference Proceedings T.I. Ören, C.M. Delfosse, C.M. Shub (Eds.)

GENERATING RANDOM VARIATES FROM A DISTRIBUTION OF PHASE TYPE

Marcel F. Neuts

and

Miriam E. Pagano Department of Mathematical Sciences University of Delaware Newark, Delaware 19711 A distribution  $F(\cdot)$  on  $[0,\infty)$  is a PH-distribution if it is that of the time until absorption in a finite state Markov process with generator

$$Q = \begin{vmatrix} T & \underline{T}^{\circ} \\ \underline{0} & 0 \end{vmatrix}$$

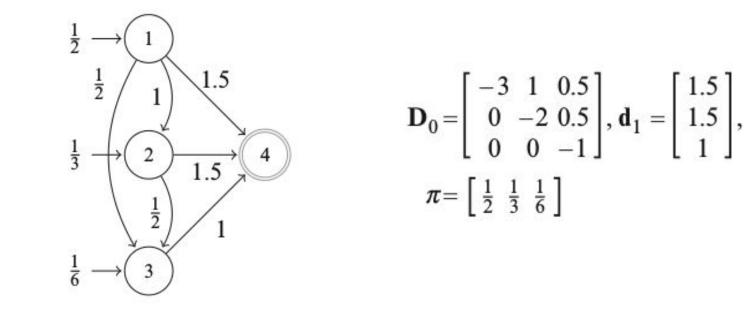
and initial probability vector  $(\underline{\alpha}, \alpha_{m+1})$ . T is a non-singular matrix of order m and satisfies  $T_{ii} < 0$ , for  $1 \le i \le m$ , and  $T_{ij} \ge 0$ , for  $i \ne j$ . Also  $T\underline{e} + \underline{T}^{\circ} = \underline{0}$  and  $\underline{\alpha} = +\alpha_{m+1} = 1$ , where  $\underline{e}$  denotes a column vector of appropriate

dimension with all components equal to one. The distribution  $F(\cdot)$  is then given by

br



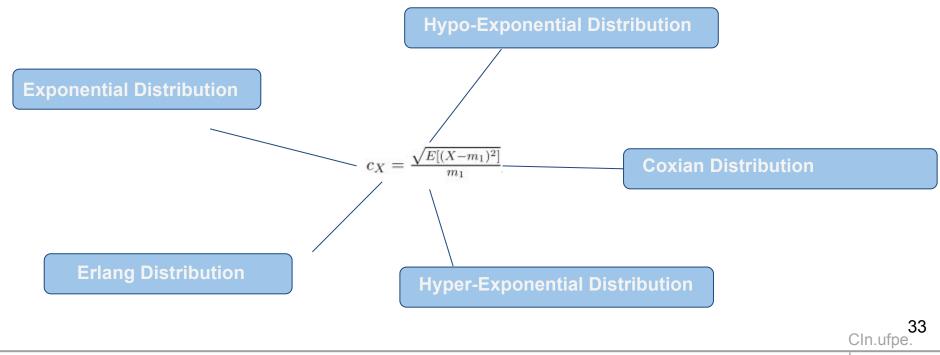
### • PHD - PHase-type Distribution



32 Cln.ufpe. br



• Can be divided into various subclasses:



## Acyclic Phase-Type Distributions

### • Exponential Distribution

- $\circ~$  Simplest case of a PHD with a single transient state.
- **Memoryless:** Prob(X > t + s | X > t) = Prob(X > s) for all  $t, s \ge 0$ .
- Mean:  $E[X] = \frac{1}{\lambda}$  Variance:  $VAR[X] = \frac{1}{\lambda^2}$

$$a \xrightarrow{1 \to (1) \to (2)} \qquad \qquad b = \begin{bmatrix} -\lambda & \lambda \\ 0 & 0 \end{bmatrix}$$

## Acyclic Phase-Type Distributions

## • Erlang Distribution

- It is a sequence of *n* exponential distributions with the same rate.
- **Denoted by:**  $E(n, \lambda)$
- Initial probability vector:  $\pi = [1, 0, ..., 0]$

$$\mathbf{a} \qquad \qquad \mathbf{b} \\ 1 \longrightarrow \underbrace{(1 \longrightarrow 2)}_{\lambda} \xrightarrow{\lambda} \underbrace{(1$$

## Acyclic Phase-Type Distributions

## Hypo-Exponential Distribution

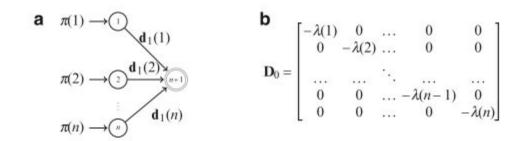
• Generalized Erlang Distribution.

• Rates  $\lambda(1), \dots, \lambda(n)$  are not necessarily identical.

# Acyclic Phase-Type Distributions

## Hyper-Exponential Distribution

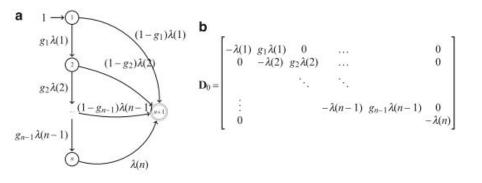
- It is a convex mixture of *n* exponential distributions.
- $\circ \pi(i) > 0$  for all phases.



# Acyclic Phase-Type Distributions

## • Coxian Distribution

- Can be considered as a mixture of hypo- and hyper-exponential distributions.
- $\circ~$  Generalized Erlang distributions with preemptive exit options.
- Initial distribution vector:  $\pi = [1, 0, ..., 0]$



38

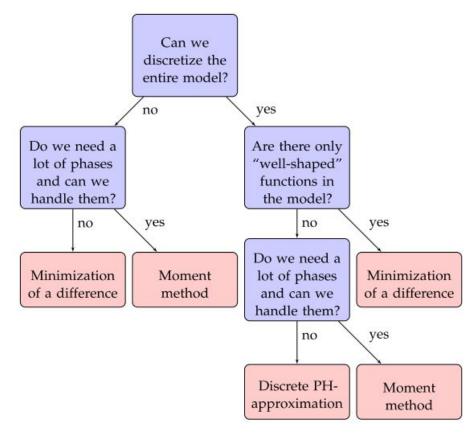


## • What is it?

- Approach for the PH-approximation
- Member of fitting techniques:
  - they utilize incomplete information of the original distribution

## • Less than 3

- Matching only the first moment m1 is simply achieved by the exponential distribution with rate  $\lambda = \frac{1}{m1}$
- In techniques matching two moments, the coefficient of variation Cv is usually used instead of the second moment, because of consequently easier representation of results.
- Distributions with 0 < Cv < 1 can be approximate by a mix of two Erlang distribuitions com  $\gamma$  e  $\gamma$ -1 phases an same rate  $\lambda$ .



41 Cln.ufpe. br

- Muitas vezes, uma atividade observada só pode ser modelada por uma variável aleatória de natureza não exponencial.
- É possível, entretanto, utilização combinações de transições exponenciais, PH-distribuitions.
- Para encontrar a distribuição, duas atividades são necessárias:
  - Determinar o tipo de aproximação necessária.
  - Encontrar os parâmetros numéricos da aproximação.
- Moments Matching algorithms fazem mapeamento de distribuições empíricas em combinações de exponenciais.

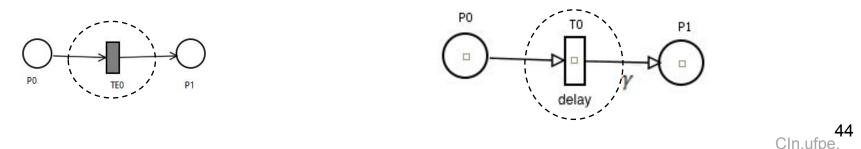
 Em SPN, uma abordagem para representar uma atividade com distribuição empírica é aproximá-la por uma distribuição exponencial, utilizando do primeiro momento.

média das durações mensuradas.

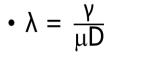
• Resultados melhores podem ser obtidos utilizando outros algoritmos que matching mais momentos.

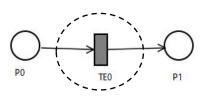
- Dado a média (μD) e o desvio padrão (σD) de uma distribuição empírica:
  - Se Cv = 1, então uma transição exponencial é suficiente, com um único parâmetro λ.

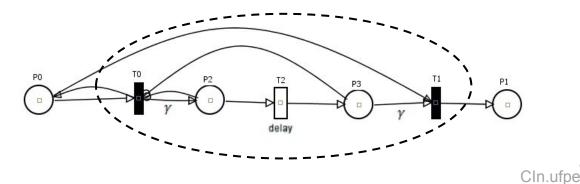
• 
$$\lambda = \frac{1}{\mu D}$$



- Dado a média (μD) e o desvio padrão (σD) de uma distribuição empírica:
  - Se Cv < 1 e Cv<sup>-1</sup> ∈ Z, então aproxima-se para uma distribuição de Erlang, e é necessário estimar dois parâmetros, λ, γ.
  - $\gamma = Cv^{-2}$







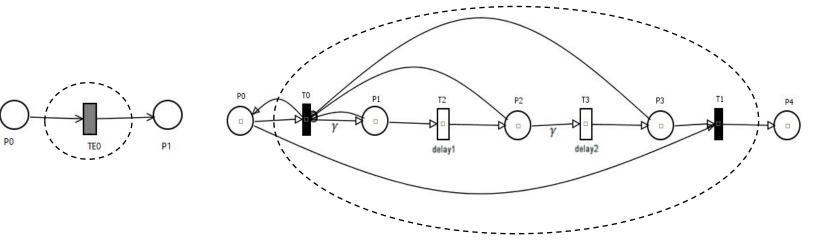
45

- Dado a média (μD) e o desvio padrão (σD) de uma distribuição empírica:
  - Se Cv < 1 e Cv<sup>-1</sup> ∉ Z, então aproxima-se para uma distribuição de Hipoexponencial e é necessário estimar três parâmetros: γ, λ<sub>1</sub> e λ<sub>2</sub>.

• 
$$Cv^{-2} -1 \le \gamma < Cv^{-2}$$
  
•  $\lambda_1 = \frac{\gamma + 1}{\mu D \pm \sqrt{\gamma(\gamma + 1)\sigma D^2 - \gamma \mu D^2}}$   
•  $\lambda_2 = \frac{\gamma + 1}{\gamma \mu D \pm \sqrt{\gamma(\gamma + 1)\sigma D^2 - \gamma \mu D^2}}$ 



• Hipoexponencial



- Dado a média (μD) e o desvio padrão (σD) de uma distribuição empírica:
  - Se Cv > 1, então aproxima-se para uma distribuição de Hiperexponencial e é necessário estimar três parâmetros: W<sub>1</sub>, W<sup>2</sup> e λ<sub>h</sub>.

48

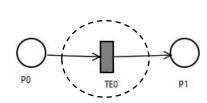
Cln.ufpe

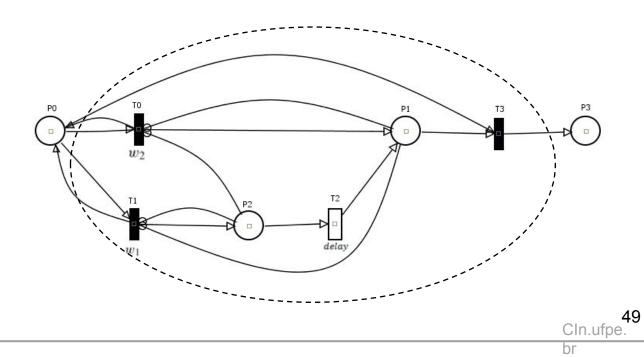
br

- $W_1 = \frac{2\mu D^2}{\mu D^2 + \sigma D^2}$
- $W_2 = 1 W_1$ •  $\lambda_h = \frac{2\mu D}{\mu D^2 + \sigma D^2}$



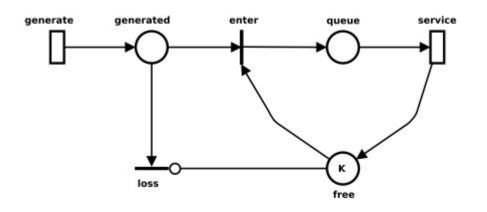
• Hiperexponencial

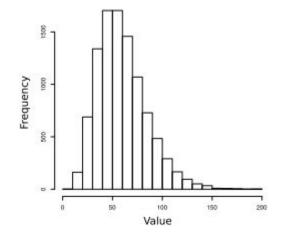






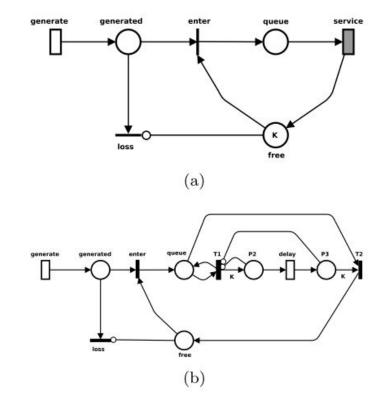
## • Service time not exponential





50 Cln.ufpe. br





- Using phase-type...
- MSL (Mercury)

Cln.ufpe.

51



#### The Application of Phase Type Distributions for Modelling Queuing Systems

Eimutis VALAKEVICIUS Department of Mathematical Research in Systems Kaunas University of Technology Kaunas, LT - 51368, Lithuania

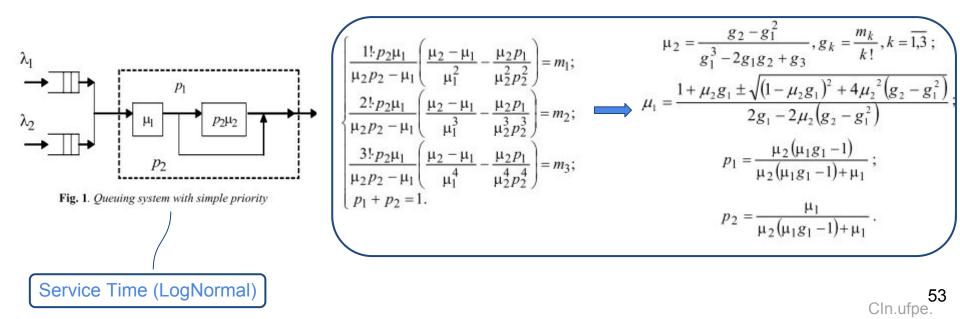
#### ABSTRACT

Queuing models are important tools for studying the performance of complex systems, but despite the substantial queuing theory literature, it is often necessary to use approximations in the case the system is nonmarkovian. Phase type distribution is by now indispensable tool in creation of queuing system models. The purpose of this paper is to suggest a method and software for evaluating queuing approximations. A numerical queuing model with priorities is used to explore the behaviour of exponential phase-type approximation of service-time distribution. The performance of queuing systems described in the event language is used for generating the set of states and transition matrix between them. Two examples of numerical models are presented - a queuing system model with priorities and a queuing system model with quality control.



### The Application of Phase Type Distributions for Modelling Queuing Systems

Eimutis VALAKEVICIUS Department of Mathematical Research in Systems Kaunas University of Technology Kaunas, LT - 51368, Lithuania





### The Application of Phase Type Distributions for Modelling Queuing Systems

Eimutis VALAKEVICIUS Department of Mathematical Research in Systems Kaunas University of Technology Kaunas, LT - 51368, Lithuania

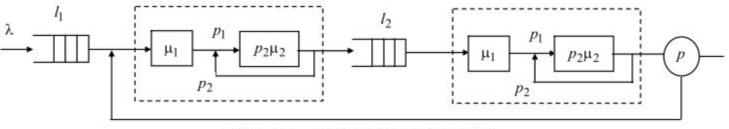


Fig.2. Queuing system with quality control



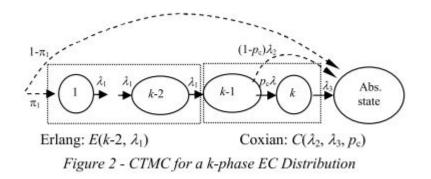
#### Haitao Liao, PhD, University of Arizona Huairui Guo, PhD, ReliaSoft Corporation

Key Words: accelerated life testing, Erlang-Coxian distribution, maximum likelihood estimation

#### SUMMARY & CONCLUSIONS

Accelerated life testing (ALT) is widely used to expedite failures of a product in a short time period for predicting the product's reliability under normal operating conditions. The resulting ALT data are often characterized by a probability distribution, such as Weibull, Lognormal, Gamma distribution, along with a life-stress relationship. However, if the selected failure time distribution is not adequate in describing the ALT data, the resulting reliability prediction would be misleading. This paper proposes a generic method that assists engineers in modeling ALT data. The method uses Erlang-Coxian (EC) distributions, which belong to a particular subset of phase-type (PH) distributions, to approximate the underlying failure time distributions arbitrarily closely. To









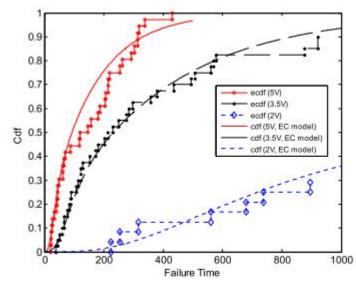
Stress level	("+"	Failure tin the unit is c				
5 volts	20.5         22.3           34.1         39.6           47.7         61.6           87.8         118           180.9         187           215.2         218           304         313           430.2         304	23.2 41.8 62.1 3 120.1 7 204 7 254.1	24.7 43.6 65.5 145.4 206.7 262.6 317.9	26 44.9 70.8 157.4 213.9 293 337.7	Weibull? LogNormal?	$\begin{array}{c} 0.9\\ 0.8\\ 0.7\\ 0.6\\ 0.6\\ 0.5\\ 0.5\\ 0.5\\ 0.5\\ 0.5\\ 0.5\\ 0.5\\ 0.5$
3.5 volts	37.8         43.6           65.9         75.6           106.6         113.1           151.8         171.           230.7         249.           358.5         379.           561.1         570           890+         890+	82.5 121.1 7 181 9 275.6 8 434.5 577.7	58.6 88.1 121.5 202.7 285 493.1 876.3 941+	65.5 89 128.3 211.7 296.2 506.4 922 941+		
2 volts	223.1 254 737 894 930.5+ 930 930.5+ 930 930.5+ 930	316.7 4 930.5- 5+ 930.5- 5+ 930.5-	560.2 - 930.5+ - 930.5+ - 930.5+ - 930.5+	679		0 200 400 600 800 1000 Failure Time Figure 3 - Empirical cdf's of the Failure Times under Different Voltage Levels Cln.ufpe.

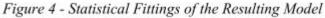
br

lature Lamps



Values of k	MLEs of parameters	Log-likelihood lnL
3: $E(1, \lambda_1)$ & $C(\lambda_2, \lambda_3, p_c)$	$\begin{array}{l} \alpha_0 = 2.9091; \\ \alpha_1 = 0.5762; \\ \lambda_1 = 0.0026; \\ \lambda_2 = 0.0026; \\ \lambda_3 = 0.0003; \\ p_c = 0.6730; \end{array}$	-518.5038
4: $E(2, \lambda_1)$ & $C(\lambda_2, \lambda_3, p_c)$	$\begin{array}{l} \alpha_0 = 2.8807; \\ \alpha_1 = 0.5730; \\ \lambda_1 = 0.0045; \\ \lambda_2 = 0.0045; \\ \lambda_3 = 0.0003; \\ p_c = 0.6980; \end{array}$	-516.4058
5: $E(3, \lambda_1)$ & $C(\lambda_2, \lambda_3, p_c)$	$\begin{aligned} \alpha_0 &= 2.8182; \\ \alpha_1 &= 0.5693; \\ \lambda_1 &= 0.0068; \\ \lambda_2 &= 0.0068; \\ \lambda_3 &= 0.0004; \\ p_c &= 0.7269; \end{aligned}$	-515.4942
6: $E(4, \lambda_1)$ & $C(\lambda_2, \lambda_3, p_c)$	$\begin{aligned} \alpha_0 &= 2.7463; \\ \alpha_1 &= 0.5747; \\ \lambda_1 &= 0.0097; \end{aligned}$	-515.0856





58 Cln.ufpe. br



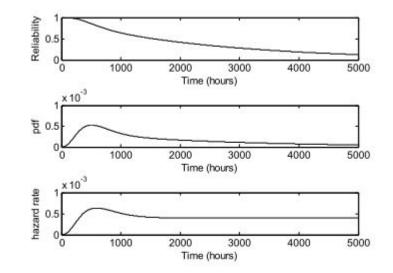


Figure 5 - Reliability Prediction using the Resulting Model (reliability function, pdf, hazard rate under 2V)





60 Cln.ufpe. br

## Definition - PH Distribution

- **R** Package actuar
  - Density Phase Type

dphtype(x, prob, rates)

- Distribution Function
  - pphtype(q, prob, rates)
- Random Generator
  - rphtype(n, prob, rates)
- Raw moments
  - mphtype(order, prob, rates)
- Moment Generator Function

mgfphtype(x, prob, rates)

61



- Jain, Raj. *The art of computer systems performance analysis*. John Wiley & Sons, 2008.
- Komárková, Zuzana. "Phase-Type Approximation Techniques." (2012).
- Buchholz, Peter. Input Modeling with Phase-Type Distributions and Markov Models: Theory and Applications. Springer, 2014.
- Pérez, Juan F., and Germán Riaño. "jPhase: an object-oriented tool for modeling phase-type distributions." *Proceeding from the 2006 workshop on Tools for solving structured Markov chains*. ACM, 2006.
- VALAKEVICIUS, Eimutis. "The Application of Phase Type Distributions for Modelling Queuing Systems." Proceedings of the second international conference simulation gaming, training and business reengineering in operations. Riga Latvia. http://www. iiisci. org. journal/CV \$/sci/pdfs/S253GBD. pdf. Accessed. Vol. 10. 2011.
- Breuer, Lothar, Dieter Baum, and D. Baum. *An Introduction to Queueing Theory: and Matrix-Analytic Methods*. Springer Science & Business Media, 2005.
- Pérez-Ocón, Rafael, and JE Ruiz Castro. "Two models for a repairable two-system with phase-type sojourn time distributions." *Reliability Engineering & System Safety* 84.3 (2004): 253-260.
- Liao, Haitao, and Huairui Guo. "A generic method for modeling accelerated life testing data." *Reliability and Maintainability Symposium* (*RAMS*), 2013 Proceedings-Annual. IEEE, 2013.



- Reinecke, Philipp, Tilman Krauß, and Katinka Wolter. "Phase-type fitting using HyperStar." *Computer Performance Engineering*. Springer Berlin Heidelberg, 2013. 164-175.
- Albuquerque Júnior, Gabriel Alves de, Paulo Romero Martins Orientador Maciel, and Ricardo Massa Ferreira Coorientador Lima.
   Modelagem e Avaliação de Desempenho Operacional e Ambiental em Cadeias de Suprimentos Verdes. Diss. Universidade Federal de Pernambuco, 2013.
- Oliveira, Danilo. "The Mercury Scripting Language Cookbook."