



**UNIVERSIDADE FEDERAL DA BAHIA  
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA  
DEPARTAMENTO DE ESTATÍSTICA**

**Bruno Damasceno de Jesus**

**Gráficos de Controle Estatístico para  
Monitoramento de Dados de Contagem com  
Superdispersão Baseados em Distribuições  
de Misturas Uniparamétricas de Poisson**

Salvador

2021

Bruno Damasceno de Jesus

**Gráficos de Controle Estatístico para  
Monitoramento de Dados de Contagem com  
Superdispersão Baseados em Distribuições  
de Misturas Uniparamétricas de Poisson**

Projeto apresentado ao Curso de graduação em Estatística, Departamento de Estatística, Instituto de Matemática e Estatística, Universidade Federal da Bahia, como requisito para aprovação na disciplina de Trabalho de Conclusão de Curso II.

**Orientador:** Paulo Henrique Ferreira da Silva

Salvador  
2021



# Resumo

A distribuição de Poisson é um modelo probabilístico discreto amplamente usado para analisar dados de contagem. Gráficos ou cartas de controle estatístico baseados nesta distribuição, como os gráficos  $c$  e  $u$ , estão bem estabelecidos na literatura. No entanto, muitos estudos sugerem a necessidade de abordagens alternativas que permitam modelar a superdispersão (variância  $>$  média), fenômeno que pode ser observado em diversas áreas, como biologia, ecologia, saúde, marketing, economia e indústria, entre outras. As distribuições de misturas uniparamétricas de Poisson, cuja literatura é extensa e importante, são capazes de modelar a variabilidade extra-Poisson, acomodando diferentes níveis de superdispersão. As distribuições pertencentes a essa classe de modelos, que inclui as distribuições Poisson-Lindley, Poisson-Shanker e Poisson-Sujatha, dentre outras, podem assim ser utilizadas como alternativas interessantes às conhecidas distribuições de Poisson e COM-Poisson (biparamétrica) para analisar dados de contagem oriundos de diferentes áreas do conhecimento. Neste trabalho, consideramos a classe de modelos probabilísticos acima mencionada (bem como os três membros citados de tal classe) para desenvolver gráficos de controle estatístico inéditos e úteis para processos de contagem, que são capazes de monitorar dados de contagem que exibem superdispersão. O desempenho dos chamados gráficos de misturas uniparamétricas de Poisson, a saber:  $PL_c$ - $PL_u$ ,  $PSh_c$ - $PSh_u$  e  $PSu_c$ - $PSu_u$ , é medido pelo comprimento médio da sequência ou ARL (do inglês *Average Run Length*), em estudos de simulação numérica. Alguns conjuntos de dados são usados para ilustrar a aplicabilidade da metodologia proposta.

**Palavras-chave:** Controle Estatístico de Processo (CEP); dados de contagem; gráfico de controle; mistura uniparamétrica de Poisson; superdispersão.

# Abstract

The Poisson distribution is a discrete model widely used to analyze count data. Statistical control charts based on this distribution, such as the  $c$  and  $u$  charts, are relatively well-established in the literature. Nevertheless, many studies suggest the need for alternative approaches that allow for modeling overdispersion, a phenomenon that can be observed in several fields, including biology, ecology, healthcare, marketing, economics, and industry. The one-parameter Poisson mixture distributions, whose literature is extensive and essential, can model extra-Poisson variability, accommodating different overdispersion levels. The distributions belonging to this class of models, including the Poisson-Lindley, Poisson-Shanker, and Poisson-Sujatha models, can thus be used as interesting alternatives to the usual Poisson and COM-Poisson distributions for analyzing count data in several areas. In this paper, we consider the class of probabilistic models mentioned above (as well as the cited three members of such a class) to develop novel and useful statistical control charts for counting processes, monitoring count data that exhibit overdispersion. The performance of the so-called one-parameter Poisson mixture charts, namely the  $PL_c$ - $PL_u$ ,  $PSh_c$ - $PSh_u$ , and  $PSu_c$ - $PSu_u$  charts, is measured by the average run length in exhaustive numerical simulations. Some data sets are used to illustrate the applicability of the proposed methodology.

**Keywords:** Statistical Process Control (SPC); count data; control chart; one-parameter Poisson mixture distribution; overdispersion.

# Lista de ilustrações

Figura 1 – Comparação de desempenho dos gráficos $PL_u$ , $u$ e $cm_pu$ , todos construídos a partir de amostras de distribuição PL (fase 2). . . . .	22
Figura 2 – Comparação de desempenho dos gráficos $PSh_c$ , $c$ e $BN-c$ , construídos a partir dos dados de erros de informação em registros de clientes. . . .	23
Figura 3 – Gráfico $PSu_c$ para o número de óbitos de COVID-19, por semana epidemiológica de notificação, no estado da Bahia, Brasil. . . . .	25

# Lista de tabelas

Tabela 1	– Probabilidades de mistura para cada modelo. . . . .	11
Tabela 2	– Caracterização de cada elemento da classe $F_1$ . . . . .	12
Tabela 3	– Caracterização de cada elemento da classe $F_2$ . . . . .	12
Tabela 4	– Expressões de $\tau_k$ ( $k = 1, 2, 3$ ) para cada elemento da classe $F_2$ . . . . .	12
Tabela 5	– Média ( $\mu$ ), variância ( $\sigma^2$ ) e índice de dispersão ( $\gamma$ ) para cada elemento da classe $F_2$ . . . . .	13
Tabela 6	– Expressão de $\theta = h(\mu)$ para cada elemento da classe $F_2$ . . . . .	14
Tabela 7	– Limites de controle para os gráficos dos tipos $c$ e $u$ com base em cada elemento da classe $F_2$ (parâmetro conhecido ou especificado). . . . .	15
Tabela 8	– Valores médios de $ARL_0$ (com desvio-padrões entre parênteses) dos gráficos $PL_c$ , $PSh_c$ e $PSu_c$ , para os diferentes cenários estudados. . . . .	18
Tabela 9	– Valores médios de $ARL_1$ (com desvios-padrões entre parênteses) do gráfico $PL_c$ , para os diferentes cenários estudados. . . . .	19
Tabela 10	– Valores médios de $ARL_1$ (com desvios-padrões entre parênteses) do gráfico $PSh_c$ , para os diferentes cenários estudados. . . . .	19
Tabela 11	– Valores médios de $ARL_1$ (com desvios-padrões entre parênteses) do gráfico $PSu_c$ , para os diferentes cenários estudados. . . . .	20
Tabela 12	– Valores médios de $ARL_0$ (com desvio-padrões entre parênteses) dos gráficos $PL_u$ , $PSh_u$ e $PSu_u$ , para os diferentes cenários estudados, . . . . .	35
Tabela 13	– Valores médios de $ARL_1$ (com desvios-padrões entre parênteses) do gráfico $PL_u$ , para os diferentes cenários estudados, . . . . .	36
Tabela 14	– Valores médios de $ARL_1$ (com desvios-padrões entre parênteses) do gráfico $PSh_u$ , para os diferentes cenários estudados, . . . . .	36
Tabela 15	– Valores médios de $ARL_1$ (com desvios-padrões entre parênteses) do gráfico $PSu_u$ , para os diferentes cenários estudados, . . . . .	37

# Sumário

Lista de ilustracões . . . . .	5
Lista de tabelas . . . . .	6
Sumario . . . . .	7
1    INTRODUCAO . . . . .	8
2    METODOLOGIA . . . . .	11
2.1    Distribuições de misturas uniparametricas de Poisson . . . . .	11
2.2    Novos graficos de controle . . . . .	13
3    ESTUDOS DE SIMULACAO . . . . .	16
3.1 $ARL_0$ . . . . .	16
3.2 $ARL_1$ . . . . .	17
4    APLICACÕES . . . . .	21
4.1    Dados artificiais . . . . .	21
4.2    Dados reais mais artificiais . . . . .	22
4.3    Dados reais . . . . .	23
5    CONCLUSÕES E PERSPECTIVAS FUTURAS . . . . .	26
Referências . . . . .	27



# 1 Introdução

Frequentemente, pesquisadores e profissionais se deparam com o desafio de analisar dados de contagem, para os quais existe a presença de superdispersão (ou sobredispersão). Na realidade, dados com essa característica são bastante comuns em diversas áreas do conhecimento, como saúde, biologia, ecologia, marketing, economia, indústria e criminologia (Hougaard et al., 1997; Okamura et al., 2012; Coly et al., 2016; Chen et al., 2020).

Estatisticamente falando, a superdispersão significa que há mais variação nos dados do que o predito pelo modelo (Chen et al., 2020). O fenômeno da superdispersão (variância  $>$  média) também pode ser interpretado como uma falha de alguns pressupostos básicos do modelo hipotético, cujas possíveis causas podem ser: a variabilidade do material experimental, correlação ou falta de independência entre as respostas dos itens individuais, amostragem por conglomerados (*clusters*) ou estrutura agrupada da população, heterogeneidade ou falta de suposição de população homogênea, variáveis omitidas ou não-observadas, pequeno tamanho da amostra, contágio, agregação e repulsão (veja, por exemplo, Kokonendji, 2014; Xekalaki, 2014, para algumas explicações detalhadas e adicionais).

Para dados com essa característica, definir uma classe de modelos paramétricos (em especial, modelos com um único parâmetro ou uniparamétricos) traz facilidades ao pesquisador, no que tange ao resgate intuitivo, em termos de média e variância. Consolidando também aspectos como interpretabilidade e explicabilidade de forma compreensível.

Apesar da distribuição (uniparamétrica) de Poisson ser largamente usada para modelar dados de contagem, ela assume equidispersão (variância = média) e, portanto, pode resultar em inferência incorreta sobre as estimativas dos parâmetros, erros padrões, testes e intervalos de confiança, quando houver presença de superdispersão (Avcı et al., 2015). Nas aplicações de Controle Estatístico de Processo (CEP), por exemplo, é sabido que um gráfico de controle baseado em um modelo de Poisson mal especificado (gráfico  $c$  ou  $u$ ), pode aumentar a taxa de alarmes falsos, ou seja, vários valores de dados (pontos ou amostras) podem ser falsamente detectados como fora de controle quando são apenas falsos positivos (Spiegelhalter, 2005; Mohammed and Laney, 2006; Albers, 2011; Saghir and Lin, 2015).

Quanto ao monitoramento estatístico de um processo através de gráficos (ou cartas) de controle, é sabido que as cartas do tipo  $c$  consideram o número total de eventos (por exemplo, defeitos) por unidade inspecionada (neste caso, o item em questão pode apresentar alguns ou mesmo vários defeitos que acabam por não prejudicar o seu bom funcionamento, como é o caso, por exemplo, de produtos eletrônicos), e as amostras utilizadas consistem do mesmo tamanho (isto é, mesmo número de unidades ou itens por amostra). A abordagem das cartas do tipo  $u$  é semelhante, mas tem a finalidade de

monitorar o número médio de defeitos por unidade inspecionada, em situações de tamanho variável das amostras utilizadas, em que geralmente trabalha-se com um tamanho médio de amostra.

Algumas alternativas interessantes e mais flexíveis para os gráficos de controle baseados na distribuição de Poisson têm sido apresentadas na literatura. Por exemplo, Sheaffer and Leavenworth (1976) e Kaminsky et al. (1992) desenvolveram gráficos de controle com base na distribuição binomial negativa (BN), bem como em seu caso particular, a distribuição geométrica. Sellers (2012) considerou a distribuição de Conway-Maxwell-Poisson (COM-Poisson) (Conway and Maxwell, 1962; Shmueli et al., 2005) para produzir gráficos de controle generalizados. Mais recentemente, Boaventura et al. (2020) propuseram novos gráficos de controle de atributos com base na distribuição de Bell (Castellares et al., 2018). Outras propostas existentes para dados de contagem dispersos consideram a mistura Poisson-gama (Cheng and Yu, 2013), a distribuição de Poisson generalizada deslocada (ou truncada em zero) (Famoye, 1994), as distribuições BerG (que é uma soma de variáveis aleatórias de Bernoulli e geométrica) e de Poisson dupla (Bourguignon et al., 2020), a distribuição Touchard (Ho et al., 2021), entre outras.

Uma abordagem popular que supera as limitações de um modelo de Poisson consiste em usar modelos de mistura finita (McLachlan and Peel, 2004), que são capazes de acomodar diferentes níveis de superdispersão (Karlis and Xekalaki, 2005). Por exemplo, a distribuição binomial negativa, que surge como um modelo de mistura de Poisson usando uma distribuição gama para a parte contínua, é uma das alternativas mais utilizadas para capturar a variabilidade extra-Poisson. Contudo, há uma extensa literatura sobre outras distribuições mistas discretas que podem permitir diferentes níveis de superdispersão. Dentre as quais, pode-se citar as seguintes distribuições uniparamétricas: Poisson-Lindley (Sankaran, 1970), Poisson-lognormal (Bulmer, 1974), Poisson-Gaussiana inversa (Shaban, 1981), Poisson-Janardan (Shanker et al., 2014), Poisson-Amarendra (Shanker, 2016a), Poisson-Shanker (Shanker, 2016b), Poisson-Sujatha (Shanker and Hagos, 2016), Poisson-Aradhana (Shanker, 2017b), Poisson-Akash (Shanker, 2017a), Poisson-Garima (Shanker, 2017c) e Poisson-Ishita (Shukla and Shanker, 2019).

Essa grande classe de modelos de probabilidade discretos, que contém as distribuições Poisson-Lindley (PL), Poisson-Shanker (PSh) e Poisson-Sujatha (PSu), motivou o desenvolvimento neste trabalho de novos gráficos de controle dos tipos  $c$  e  $u$ , para monitorar os números total e médio de eventos por unidade de inspeção, respectivamente. Os chamados gráficos de misturas uniparamétricas de Poisson representam alternativas aos gráficos de controle mencionados anteriormente, para o monitoramento de processos de contagem com dados superdispersos. Em particular, este trabalho focar-se-á nos três modelos (citados acima) desta relevante classe de modelos, para derivar os chamados gráficos (ou cartas) de controle  $PL_c$ - $PL_u$ ,  $PSh_c$ - $PSh_u$ , e  $PSu_c$ - $PSu_u$ . Pela maneira como são construídos, tais gráficos têm como principal finalidade identificar mudanças destacadas no

parâmetro (único) do processo.

Vale ressaltar que, em um gráfico de controle de Shewhart, que é o caso das ferramentas CEP propostas neste trabalho, o objetivo principal é detectar mudanças maiores (isto é, de maior magnitude) nos parâmetros do processo. Uma vez que esse tipo de gráfico de controle ignora dados históricos, ele é comumente referido como “sem memória”.

O restante deste trabalho está organizado da seguinte maneira. No Capítulo 2, especificamente na Seção 2.1, é revista a classe de distribuições de misturas uniparamétricas de Poisson e algumas de suas propriedades básicas. Enquanto que, na Seção 2.2, são apresentados os novos gráficos de controle baseados nessa classe de modelos discretos, particularmente, para três de seus membros (PL, PSh e PSu). No Capítulo 3, são apresentados os estudos de simulação numérica realizados para avaliar o desempenho dos gráficos de controle usando essas três distribuições. No Capítulo 4, é ilustrada a aplicabilidade e a utilidade dos gráficos de controle propostos, por meio de alguns conjuntos de dados. Finalmente, o trabalho é concluído com algumas observações finais e orientações para desenvolvimentos futuros no Capítulo 5.

## 2 Metodologia

### 2.1 Distribuições de misturas uniparamétricas de Poisson

Nesta seção, é realizada uma breve revisão da classe de distribuições de misturas uniparamétricas de Poisson e algumas de suas propriedades básicas. Particularmente, para três membros dessa classe de modelos discretos, nomeadamente as distribuições PL, PSh e PSu. É discutida também a estimação do único parâmetro por meio do método de máxima verossimilhança.

A classe de interesse contém distribuições que surgem quando se assume que o parâmetro de taxa  $\lambda$  de uma distribuição de Poisson se comporta de acordo com uma variável aleatória no espaço paramétrico  $\Lambda = \mathbb{R}_+$ . Neste caso, pode-se olhar para especificações que provêm flexibilidade ao descrever o parâmetro  $\lambda$  e que levam, por exemplo, a uma distribuição uniparamétrica como produto final. Os modelos resultantes podem ser vistos como alternativas interessantes à distribuição de Poisson usual, na tarefa de analisar dados de contagem na presença de heterogeneidade (superdispersão).

Como escolhas possíveis para descrever o comportamento do parâmetro de taxa da distribuição de Poisson, pode-se considerar as distribuições uniparamétricas de Lindley (L; Lindley, 1958), Shanker (Sh; Shanker, 2015) e Sujatha (Su; Shanker, 2016c). Elas compartilham a característica de ter a parte contínua da composição sendo definida pela mistura finita de distribuições gama, mudando apenas as probabilidades de mistura para cada modelo, conforme exibido na Tabela 1.

Tabela 1 – Probabilidades de mistura para cada modelo.

Modelo	Gamma(1, $\theta$ )	Gamma(2, $\theta$ )	Gamma(3, $\theta$ )
L	$\theta(\theta + 1)^{-1}$	$(\theta + 1)^{-1}$	-
Sh	$\theta^2(\theta^2 + 1)^{-1}$	$(\theta^2 + 1)^{-1}$	-
Su	$\theta^2(\theta^2 + \theta + 2)^{-1}$	$\theta(\theta^2 + \theta + 2)^{-1}$	$2(\theta^2 + \theta + 2)^{-1}$

Essas distribuições provêm flexibilidade em relação à sua forma, que é controlada por um parâmetro único  $\theta \in \Theta = \mathbb{R}_+$ . Seja  $F_1 = fg(., \theta) : \theta \in \Theta$  uma classe de funções densidade de probabilidade que, para uma variável aleatória positiva contínua  $X$ , tem elementos da forma:

$$g(x; \theta) = \frac{m_1(x, \theta)}{c_1(\theta)} e^{-\theta x}, \quad x \in \mathbb{R}_+,$$

em que  $c_1(\theta) > 0$  é a constante normalizadora e  $\theta$  é o parâmetro de forma. Dado a abrangência e a flexibilidade da classe  $F_1$ , as funções  $m_1(x, \theta)$  e  $c_1(\theta)$ , para as três distribuições consideradas aqui, são apresentadas na Tabela 2.

Tabela 2 – Caracterização de cada elemento da classe  $F_1$ .

Modelo	$c_1(\theta)$	$m_1(x, \theta)$
L	$(\theta + 1)\theta^{-2}$	$x + 1$
Sh	$(\theta^2 + 1)\theta^{-2}$	$x + \theta$
Su	$(\theta^2 + \theta + 2)\theta^{-3}$	$x^2 + x + 1$

Uma classe  $F_2$  de modelos probabilísticos discretos surge quando um elemento de  $F_1$  é escolhido para descrever o parâmetro de taxa  $\lambda$  da distribuição de Poisson. Assim, seja  $F_2 = \{f(\cdot; \theta) : \theta \in \Theta\}$  uma classe cujos elementos são distribuições de misturas uniparamétricas de Poisson obtidas de  $g(\cdot; \theta) \in F_1$ . Ou seja, uma variável aleatória discreta  $Y$  é distribuída de acordo com um elemento de  $F_2$  se  $Y|\lambda \sim \text{Poisson}(\lambda)$  e  $\lambda \sim g(\cdot; \theta) \in F_1$ . Então, a distribuição incondicional de  $Y$ , denotada como  $f(y; \theta)$ , é dada por:

$$f(y; \theta) = \frac{m_2(y, \theta)}{c_2(\theta)}, \quad y \in \mathbb{N}_0. \quad (2.1)$$

A Tabela 3 caracteriza os elementos escolhidos da classe  $F_2$  (ou seja, as distribuições PL, PSh e PSu) em termos das funções acima.

Tabela 3 – Caracterização de cada elemento da classe  $F_2$ .

Modelo	$c_2(\theta)$	$m_2(y, \theta)$
PL	$(\theta + 1)^3\theta^{-2}$	$(\theta + y + 2)(\theta + 1)^{-y}$
PSh	$(\theta + 1)^2(\theta^2 + 1)\theta^{-2}$	$(\theta^2 + \theta + y + 1)(\theta + 1)^{-y}$
PSu	$(\theta + 1)^3(\theta^2 + \theta + 2)\theta^{-3}$	$[y^2 + y(\theta + 4) + (\theta^2 + 3\theta + 4)](\theta + 1)^{-y}$

Se  $Y \sim f(\cdot; \theta) \in F_2$ , então seu momento fatorial de ordem  $r$  sobre a origem é dado por:

$$\mu'_r = \frac{r!(\theta + \tau_1 + \tau_2)}{\theta^r(\theta + \tau_3)}, \quad (2.2)$$

em que as expressões para  $\tau_k$  ( $k = 1, 2, 3$ ), dependendo dos elementos de  $F_2$ , são mostradas na Tabela 4.

Tabela 4 – Expressões de  $\tau_k$  ( $k = 1, 2, 3$ ) para cada elemento da classe  $F_2$ .

Modelo	$\tau_1$	$\tau_2$	$\tau_3$
PL	$r$	$1$	$1$
PSh	$r\theta^{-1}$	$\theta^{-1}$	$\theta^{-1}$
PSu	$r + 1$	$(r + 1)(r + 2)\theta^{-1}$	$1 + 2\theta^{-1}$

Da Equação (2.2), pode-se obter a média ( $\mu$ ), a variância ( $\sigma^2$ ) e o índice de dispersão ( $\gamma = \sigma^2/\mu$ ) das distribuições PL, PSh e PSu. Essas quantidades são apresentadas na

Tabela 5. Observe que, qualquer que seja  $\theta \geq \Theta$ , tem-se que  $\gamma > 1$ , o que implica, numa substituição direta, que a variância é maior do que a média ( $\sigma^2 > \mu$ ), logo, todas as distribuições podem acomodar superdispersão. Por outro lado, para  $\theta \rightarrow 1$ , tem-se que  $\gamma \rightarrow 1$ , o que significa que os elementos da classe  $F_2$  têm a propriedade de equidispersão (distribuição de Poisson) para grandes valores de  $\theta$ , ou equivalentemente, para pequenos valores de  $\mu$ .

Tabela 5 – Média ( $\mu$ ), variância ( $\sigma^2$ ) e índice de dispersão ( $\gamma$ ) para cada elemento da classe  $F_2$ .

Modelo	$\mu$	$\sigma^2$	$\gamma$
PL	$\frac{\theta + 2}{\theta(\theta + 1)}$	$\frac{\theta^3 + 4\theta^2 + 6\theta + 2}{\theta^2(\theta + 1)^2}$	$1 + \frac{\theta^2 + 4\theta + 2}{\theta(\theta + 1)(\theta + 2)}$
PSh	$\frac{\theta^2 + 2}{\theta^2(\theta + 1)}$	$\frac{\theta^5 + \theta^4 + 3\theta^3 + 4\theta^2 + 2\theta + 2}{\theta^2(\theta^2 + 1)^2}$	$1 + \frac{\theta^4 + 4\theta^2 + 2}{\theta(\theta^2 + 1)(\theta^2 + 2)}$
PSu	$\frac{\theta^2 + 2\theta + 6}{\theta(\theta^2 + \theta + 2)}$	$\frac{\theta^5 + 4\theta^4 + 14\theta^3 + 28\theta^2 + 24\theta + 12}{\theta^2(\theta^2 + \theta + 2)^2}$	$1 + \frac{\theta^4 + 4\theta^3 + 18\theta^2 + 12\theta + 12}{\theta(\theta^2 + \theta + 2)(\theta^2 + 2\theta + 6)}$

Em relação à geração de números pseudo-aleatórios de  $Y = f(\cdot; \theta) \geq F_2$ , pode-se facilmente recorrer à representação de mistura finita exibida na Tabela 1 para, primeiro, gerar um valor  $\lambda = g(\cdot; \theta) \geq F_1$  e, em seguida, produzir  $yj\lambda \sim \text{Poisson}(\lambda)$ . Esse procedimento pode ser repetido  $n$  vezes para obter uma amostra aleatória de tamanho  $n$  da distribuição discreta desejada.

O parâmetro  $\theta$  pode ser estimado pelo método de máxima verossimilhança (MV). Contudo, o estimador de MV  $\hat{\theta}$  de  $\theta$  não tem expressão analítica explícita para qualquer distribuição pertencente à classe  $F_2$ . Então, deve-se recorrer a métodos numéricos, como o algoritmo Broyden-Fletcher-Goldfarb-Shanno de memória limitada (L-BFGS-B) (para maiores detalhes, ver, por exemplo, Nocedal and Wright, 2006)<sup>1</sup>, que está implementado na função `optim` do *software* R (R Core Team, 2018), para obter a estimativa de MV.

Uma reparametrização alternativa e útil (por exemplo, em uma estrutura de regressão) para  $Y = f(\cdot; \theta) \geq F_2$  consiste em escrever  $\theta$  em termos da média da distribuição  $\mu = E[Y]$ , ou seja,  $\theta = h(\mu)$ . A Tabela 6 mostra a função  $h(\mu)$  para as distribuições PL, PSh e PSu.

## 2.2 Novos gráficos de controle

Vamos supor que um processo (um processo de fabricação ou de negócios, por exemplo) gera eventos (por exemplo, defeitos ou não-conformidades) de acordo com qualquer distribuição pertencente à classe  $F_2$ . Ou seja, se  $Y$  denota o número de eventos por unidade de inspeção, então a função massa de probabilidade de  $Y$  é dada pela Equação

<sup>1</sup> O algoritmo L-BFGS-B pode lidar adequadamente com restrições de caixa simples em variáveis, que é o caso aqui, uma vez que  $\theta > 0$ .

Tabela 6 – Expressão de  $\theta = h(\mu)$  para cada elemento da classe  $F_2$ .

Modelo	$h(\mu)$
PL	$\frac{(\mu - 1) \sqrt{(\mu - 1)^2 + 8\mu}}{2\mu}$
PSH	$\frac{1}{3\mu} \left[ 1 + \frac{s(\mu)}{2} \frac{2(3\mu^2 - 1)}{s(\mu)} \right]$ , com $s(\mu) = \left[ 12\mu \sqrt{3(4\mu^4 + 71\mu^2 + 8)} + 180\mu^2 + 8 \right]^{1/3}$
PSu	$\log \left( \frac{[p(u) + q(u)]^{1/3}}{\mu} \frac{(\mu - 1)(5\mu + 1)}{\mu[p(u) + q(u)]^{1/3}} \frac{\mu - 1}{\mu} \right) \log(3)$ , com $p(\mu) = 3\mu \sqrt{3(7\mu^4 + 28\mu^3 + 171\mu^2 + 32\mu + 5)}$ e $q(\mu) = 8\mu^3 + 66\mu^2 + 6\mu + 1$

(2.1). Suponha também que  $Y_1, Y_2, \dots, Y_n$  seja uma amostra aleatória de tamanho  $n$  de  $Y = f(\cdot; \theta) \geq F_2$ , e considere as estatísticas  $T = \sum_{i=1}^n Y_i$  e  $\bar{Y} = T/n = \sum_{i=1}^n Y_i/n$ , que representam o número total de eventos e o número médio de eventos por unidade de inspeção, respectivamente. Usando as propriedades dos operadores esperança e variância, tem-se que:

$$E[T] = nE[Y], \quad \text{Var}[T] = n\text{Var}[Y], \quad E[\bar{Y}] = E[Y] \quad \text{e} \quad \text{Var}[\bar{Y}] = \frac{\text{Var}[Y]}{n}.$$

Então, a partir dos resultados acima, usando as medidas teóricas (especificamente, a média e variância) em relação a cada elemento da classe  $F_2$ , pode-se facilmente derivar novos gráficos de controle com base nessa classe proposta de distribuições discretas, da maneira usual para gráficos de controle de Shewhart. Particularmente, com base nas distribuições PL, PSh e PSu, cujas médias e variâncias estão apresentadas na Tabela 5. Ou seja, será usada a regra geral “Média  $\pm L$  DP” para cálculo da linha central (LC = Média) e dos limites superior e inferior de controle (LSC = Média +  $L$  DP e LIC = Média  $- L$  DP, respectivamente) dos gráficos propostos, em que  $L$  denota a largura dos limites de controle (no programa de controle de qualidade Seis Sigma tradicional,  $L = 3$ ) e DP representa o desvio-padrão (que é a raiz quadrada da variância).

Portanto, assumindo que o verdadeiro valor de  $\theta$  é conhecido (ou um valor de referência é fornecido), os limites de controle para os gráficos dos tipos  $c$  e  $u$  com base nas distribuições PL, PSh e PSu são mostrados na Tabela 7. Vale ressaltar que, uma vez que as distribuições de misturas uniparamétricas de Poisson se aproximam da distribuição de Poisson para grandes valores de  $\theta$ , os limites de controle apresentados na Tabela 7 também se aproximam dos limites de controle para os casos especiais dos gráficos  $c$  e  $u$  derivados da distribuição de Poisson. Além disso, os gráficos do tipo  $c$  devem ser usados para monitorar o número total de eventos por unidade de inspeção, e não aplicados nos casos em que os tamanhos das amostras são desiguais, pois tanto a LC quanto os limites de controle irão variar com o tamanho da amostra. Assim, os gráficos do tipo  $u$  são re-

comendados para essas situações de amostras com tamanho variável, pois são mais fáceis de interpretar (a LC dos gráficos do tipo  $u$  não irá variar entre as amostras).

Tabela 7 – Limites de controle para os gráficos dos tipos  $c$  e  $u$  com base em cada elemento da classe  $F_2$  (parâmetro conhecido ou especificado).

Modelo	Gráfico do tipo $c$		Gráfico do tipo $u$	
PL	$\frac{n(\theta + 2)}{\theta(\theta + 1)}$	$L\sqrt{\frac{n(\theta^3 + 4\theta^2 + 6\theta + 2)}{\theta^2(\theta + 1)^2}}$	$\frac{\theta + 2}{\theta(\theta + 1)}$	$L\sqrt{\frac{\theta^3 + 4\theta^2 + 6\theta + 2}{n\theta^2(\theta + 1)^2}}$
PSH	$\frac{n(\theta^2 + 2)}{\theta^2(\theta + 1)}$	$L\sqrt{\frac{n(\theta^5 + \theta^4 + 3\theta^3 + 4\theta^2 + 2\theta + 2)}{\theta^2(\theta^2 + 1)^2}}$	$\frac{\theta^2 + 2}{\theta^2(\theta + 1)}$	$L\sqrt{\frac{\theta^5 + \theta^4 + 3\theta^3 + 4\theta^2 + 2\theta + 2}{n\theta^2(\theta^2 + 1)^2}}$
PSu	$\frac{n(\theta^2 + 2\theta + 6)}{\theta(\theta^2 + \theta + 2)}$	$L\sqrt{\frac{n(\theta^5 + 4\theta^4 + 14\theta^3 + 28\theta^2 + 24\theta + 12)}{\theta^2(\theta^2 + \theta + 2)^2}}$	$\frac{\theta^2 + 2\theta + 6}{\theta(\theta^2 + \theta + 2)}$	$L\sqrt{\frac{\theta^5 + 4\theta^4 + 14\theta^3 + 28\theta^2 + 24\theta + 12}{n\theta^2(\theta^2 + \theta + 2)^2}}$

Nos casos em que o parâmetro  $\theta$  é desconhecido ou não-especificado, pode-se estimá-lo usando todas as  $m$  amostras disponíveis, por exemplo, através do método de estimação por MV. Então, pode-se substituir  $\theta$  por sua estimativa de MV  $\hat{\theta}$  nas expressões dos limites de controle mostradas na Tabela 7, obtendo os chamados “limites de controle tentativos” (Montgomery, 2013).

É importante notar que o parâmetro  $\theta$  não tem interpretação direta, então torna-se difícil para profissionais de negócios e de indústrias, entre outros, fornecer um valor de referência ou especificação para ele. Para contornar isso, pode-se usar uma reparametrização conveniente das distribuições da classe  $F_2$  em termos da média, ou seja, considerando  $\theta = h(\mu)$  (ver Tabela 6). Quando nenhuma especificação existe, o parâmetro  $\mu$  pode ser estimado simplesmente pela média amostral geral, ou seja,

$$\hat{\mu} = \bar{\bar{Y}} = \begin{cases} (mn)^{-1} \sum_{j=1}^m \sum_{i=1}^n Y_{ij}, & \text{para os gráficos dos tipos } c \text{ e } u \text{ com tamanhos} \\ & \text{de amostra iguais,} \\ \left(\sum_{j=1}^m n_j\right)^{-1} \sum_{j=1}^m \sum_{i=1}^{n_j} Y_{ij}, & \text{para os gráficos do tipo } u \text{ com tamanhos de} \\ & \text{amostra desiguais,} \end{cases}$$

em que  $Y_{ij}$  representa o número de eventos para a  $i$ -ésima unidade de inspeção da  $j$ -ésima amostra. Ambos os estimadores podem ser facilmente comprovados como não-viesados para  $\mu$ . Mas, uma vez que, neste caso, os limites de controle dos gráficos do tipo  $u$  também irão variar com o tamanho da amostra, pode-se substituir  $n$  por  $n_j$  em suas expressões, ou alternativamente, conforme sugerido por Montgomery (2013), basear os cálculos dos limites de controle em um tamanho amostral médio  $\bar{n}$ , que resulta em limites constantes e é particularmente útil se os gráficos forem apresentados à gerência.

Devido à sua vantagem prática e simplicidade, consideraremos a seguir os limites de controle expressos em termos de  $\mu$ .



## 3 Estudos de simulacao

Neste capítulo são apresentados estudos de simulação de Monte Carlo (MC) para avaliar a performance dos gráficos de controle baseados nas distribuições PL, PSh e PSu. Todos os procedimentos computacionais foram implementados usando o *software* estatístico R versão 4.0.2 (R Core Team, 2020). Os códigos R usados para gerar os gráficos de controle propostos com tamanhos de amostra iguais e desiguais estão disponíveis no Apêndice A.

O comprimento médio da sequência (ARL, do inglês *Average Run Length*) é uma métrica largamente usada por pesquisadores para avaliar o desempenho dos gráficos de controle. Os ditos ARL sob controle e ARL fora de controle são denotados por  $ARL_0$  e  $ARL_1$ , respectivamente.

Seja  $Y$  o resultado ou saída de um processo de contagem que segue uma distribuição de mistura uniparamétrica de Poisson reparametrizada em termos de sua média, ou seja,  $Y \sim f(\cdot; \mu) \geq F_2$ . Além disso, seja  $\mu_s$  o parâmetro média de não-conformidades deslocado após ocorrer uma mudança em  $\mu$ , isto é,  $Y \sim f(\cdot; \mu_s) \geq F_2$ . Para os gráficos de controle propostos com base nessa classe de modelos discretos, o  $ARL_0$  e o  $ARL_1$  são definidos como segue (Saghir and Lin, 2015):

$$ARL_0 = \frac{1}{\alpha}, \quad \text{para } \alpha = 1 - P(\text{LIC} < Y < \text{LSC} | \mu),$$

$$ARL_1 = \frac{1}{1 - \beta}, \quad \text{para } \beta = P(\text{LIC} < Y < \text{LSC} | \mu_s).$$

No programa de controle de qualidade Seis Sigma tradicional,  $\alpha = 0,0027$  e, portanto,  $ARL_0 = 1/0,0027 \approx 370$ . Isso significa que, embora o processo esteja sob controle, um sinal de fora de controle (ou alarme falso) será gerado a cada 370 amostras, em média (Montgomery, 2013). Por outro lado, valores de  $ARL_1 \approx 1$  são desejados, especialmente para mudanças de grande magnitude na média do processo (gráficos de controle de Shewhart).

### 3.1 $ARL_0$

Sem perda de generalidade, nesta seção são considerados processos PL, PSh e PSu com taxas médias de não-conformidades:  $\mu = 3$  e 15 (que também correspondem, respectivamente, a menores - aproximando-se da distribuição de Poisson - e maiores níveis de superdispersão), bem como três valores distintos para a probabilidade de alarme falso:  $\alpha = 0,1, 0,01$  e 0,0027 (que correspondem, respectivamente, a  $ARL_0 = 10, 100$  e 370). Também são considerados diferentes tamanhos de amostra para cada processo:  $n = 50, 100$  e 500 (ou  $\bar{n} = 50, 100$  e 500, no caso dos gráficos do tipo  $u$  com tamanho de amostra

variável, em que cada tamanho amostral  $n_j$  é obtido por  $n_j = \bar{n} + k_j$ , para  $j = 1, \dots, m$ , sendo  $k_j$  uma constante inteira variando no intervalo  $[\bar{n}/5; \bar{n}/5]$ , além de diferentes quantidades de amostras na fase 1:  $m = 20, 50$  e  $100$ .

As Tabelas 8 e 12 (ver Apêndice B) mostram os resultados obtidos de 1.000 simulações de MC com  $m = 5.000$  amostras de fase 2 cada <sup>1</sup>, executadas para cada cenário estudado (isto é, variando o número de amostras da fase 1, o tamanho amostral, a distribuição de mistura uniparamétrica de Poisson, o parâmetro da distribuição e a probabilidade de alarme falso). Particularmente, os resultados de  $ARL_0$  na Tabela 8 dizem respeito aos gráficos do tipo  $c$  (tamanhos de amostra iguais), enquanto a Tabela 12 contém os valores de  $ARL_0$  dos gráficos do tipo  $u$  (tamanhos de amostra desiguais), com base nas distribuições PL, PSh e PSu.

Apesar de terem sido observadas, para alguns casos, diferenças leves a moderadas entre as médias das 1,000 estimativas de  $ARL_0$  e seus valores-alvos, o que é de fato esperado devido ao efeito da estimação de parâmetros sobre as propriedades dos gráficos de controle (vide, por exemplo, Jensen et al., 2006; Moraes et al., 2014; Paroissin et al., 2016), os resultados exibidos nas Tabelas 8 e 12 parecem indicar um desempenho adequado dos gráficos de misturas uniparamétricas de Poisson propostos, independentemente do cenário em que foram examinados.

## 3.2 $ARL_1$

Nesta seção, é investigada a capacidade de detecção de deslocamento (ou descontrole) dos gráficos PL, PSh e PSu em termos de  $ARL_1$ , para os mesmos cenários anteriores (com a única exceção de  $\alpha = 0,0027$ , devido principalmente a limitações de espaço e tempo). São consideradas mudanças em diferentes níveis, representando aumentos percentuais  $p$  na taxa de não-conformidades do processo  $\mu$ . Os níveis assumidos são:  $p = 0,5\%$  ( $\mu_s = 3,015$  e  $15,075$ ),  $1\%$  ( $\mu_s = 3,03$  e  $15,15$ ) e  $10\%$  ( $\mu_s = 3,3$  e  $16,5$ ).

Os resultados de  $ARL_1$  obtidos das 1.000 simulações de MC com  $m = 5.000$  amostras de fase 2 cada, são apresentados nas Tabelas 9-11 e 13-15 (ver Apêndice B), respectivamente, para os gráficos dos tipos  $c$  e  $u$  com base nas três distribuições de misturas uniparamétricas de Poisson. A partir dessas tabelas, pode-se notar que, em geral, o  $ARL_1$  diminui à medida que  $p$  aumenta (apesar de algumas flutuações nos valores da medida para  $p = 0,5\%$  e  $p = 1\%$ , uma vez que os gráficos de controle de Shewhart são mais eficientes na identificação de grandes mudanças na média do processo; ver, por exemplo,

<sup>1</sup> Conforme discutido em Montgomery (2013), na fase 1 da análise de gráficos de controle (análise retrospectiva), um conjunto de dados é coletado do processo e analisado de uma vez, obtendo limites de controle tentativos para estabelecer se o processo estava sob controle quando as primeiras  $m$  amostras foram produzidas. Por outro lado, na fase 2 da análise (análise prospectiva ou monitoramento do processo), o gráfico de controle construído a partir de um conjunto de dados “limpo”, isto é, exibindo controle estatístico (limites de controle confiáveis), é usado para monitorar a produção futura.

Tabela 8 – Valores médios de  $ARL_0$  (com desvio-padrões entre parênteses) dos gráficos  $PL_c$ ,  $PSh_c$  e  $PSu_c$ , para os diferentes cenários estudados.

		$n = 50$			$n = 100$			$n = 500$			
		$PL_c$	$PSh_c$	$PSu_c$	$PL_c$	$PSh_c$	$PSu_c$	$PL_c$	$PSh_c$	$PSu_c$	
$m = 20$	$\mu = 3$	$\alpha = 0,1$	9,8713 (0,1201)	9,7542 (0,1829)	8,9543 (0,3432)	9,8976 (0,3141)	9,8143 (0,3401)	9,9744 (0,4616)	9,8988 (0,3754)	9,4893 (0,2601)	8,0848 (0,4064)
		$\alpha = 0,01$	117,5221 (0,5187)	116,8526 (0,1799)	110,8755 (0,2666)	96,6901 (0,3678)	79,2563 (0,3691)	60,4746 (0,2310)	89,2569 (0,1439)	99,2568 (0,5007)	100,4312 (0,3556)
		$\alpha = 0,0027$	383,3617 (0,7801)	364,5330 (0,6358)	315,7960 (0,8354)	372,5912 (0,6512)	361,2976 (0,9031)	303,0212 (0,8231)	385,0210 (0,7230)	380,9212 (0,9167)	377,2352 (0,7846)
	$\mu = 15$	$\alpha = 0,1$	9,2136 (0,0828)	10,2589 (0,0910)	9,3569 (0,1070)	9,6018 (0,4317)	9,2705 (0,2958)	9,9830 (0,4714)	9,4582 (0,3476)	9,9765 (0,1660)	8,7326 (0,3602)
		$\alpha = 0,01$	110,5350 (0,3327)	117,1031 (0,2392)	98,6129 (0,3169)	103,9841 (0,4553)	106,7143 (0,3221)	92,6303 (0,1734)	101,0078 (0,4701)	101,0745 (0,2504)	93,0365 (0,2195)
		$\alpha = 0,0027$	387,2320 (0,7258)	380,4840 (0,5684)	360,7604 (0,4299)	385,2600 (0,9294)	372,5051 (0,6213)	367,2088 (0,7825)	384,1753 (0,4513)	371,9852 (0,6258)	352,2568 (0,7795)
$m = 50$	$\mu = 3$	$\alpha = 0,1$	10,1411 (0,5183)	9,7548 (0,0887)	10,3416 (0,3617)	9,6014 (0,2932)	10,1759 (0,2432)	9,4101 (0,4859)	9,9005 (0,1149)	10,1615 (0,4863)	9,8737 (0,1666)
		$\alpha = 0,01$	125,3985 (0,1540)	83,0353 (0,1642)	94,0075 (0,4159)	104,5516 (0,3512)	95,3301 (0,2237)	107,2712 (0,4355)	88,5062 (0,2347)	99,1158 (0,4736)	100,4312 (0,3470)
		$\alpha = 0,0027$	389,1356 (0,6147)	371,0698 (0,9352)	358,9821 (0,4007)	385,0258 (0,7962)	374,7315 (0,7180)	363,2589 (0,6381)	371,6543 (0,8584)	375,9614 (0,3645)	370,5892 (0,9344)
	$\mu = 15$	$\alpha = 0,1$	8,9721 (0,1350)	10,1608 (0,4929)	10,0624 (0,1231)	10,1197 (0,4779)	10,1535 (0,2072)	10,0895 (0,0926)	9,8645 (0,4422)	9,8730 (0,2450)	9,7505 (0,0915)
		$\alpha = 0,01$	90,6728 (0,4705)	97,4444 (0,4529)	98,3555 (0,4482)	99,5695 (0,3745)	99,3258 (0,3693)	103,3752 (0,2048)	100,9088 (0,0938)	103,0717 (0,2731)	93,0365 (0,2435)
		$\alpha = 0,0027$	378,5896 (0,8555)	382,6371 (0,7966)	370,5896 (0,7355)	383,5141 (0,9491)	381,6973 (0,9327)	352,5825 (0,7237)	385,9863 (0,8577)	384,6958 (0,7619)	352,1478 (0,5396)
$m = 100$	$\mu = 3$	$\alpha = 0,1$	10,1427 (0,4118)	9,6433 (0,4261)	10,2796 (0,2555)	10,2844 (0,2125)	9,5468 (0,3246)	9,1944 (0,1994)	9,7259 (0,4626)	10,0231 (0,2665)	10,0605 (0,1134)
		$\alpha = 0,01$	99,1840 (0,3987)	70,5591 (0,3534)	92,7695 (0,1370)	105,1963 (0,3353)	90,9086 (0,2867)	90,5216 (0,4324)	100,7233 (0,0920)	101,2104 (0,1698)	104,3881 (0,4854)
		$\alpha = 0,0027$	381,5824 (0,9543)	372,6589 (0,8414)	365,7896 (0,9497)	389,2154 (0,7664)	382,6984 (0,5813)	380,6589 (0,6167)	384,5698 (0,7811)	383,6985 (0,8233)	383,0012 (0,7345)
	$\mu = 15$	$\alpha = 0,1$	10,0976 (0,3508)	9,4268 (0,1949)	10,0544 (0,0998)	9,8560 (0,1309)	9,9398 (0,4037)	10,1417 (0,5119)	10,2590 (0,2231)	9,9593 (0,1161)	10,0607 (0,2280)
		$\alpha = 0,01$	96,6388 (0,5113)	76,0799 (0,5173)	102,1657 (0,3838)	110,6051 (0,1133)	109,5089 (0,4601)	98,0793 (0,3006)	103,3221 (0,2600)	102,8692 (0,2176)	102,1393 (0,1924)
		$\alpha = 0,0027$	382,5687 (0,7052)	370,9635 (0,6267)	368,1358 (0,6360)	385,6985 (0,8103)	380,6924 (0,6996)	360,2589 (0,6277)	390,0125 (0,8245)	387,3256 (0,8024)	370,9854 (0,8328)

Montgomery, 2013). Observe também que, para os gráficos de controle baseados nas distribuições PL, PSh e PSu, os valores de  $ARL_1$  são bem próximos de 1 quando  $p = 10\%$  e  $n = 500$  (ou  $\bar{n} = 500$ ).

Tabela 9 – Valores médios de  $ARL_1$  (com desvios-padrões entre parênteses) do gráfico  $PL_c$ , para os diferentes cenários estudados.

		$n = 50$			$n = 100$			$n = 500$				
		$p = 0,5\%$	$p = 1\%$	$p = 10\%$	$p = 0,5\%$	$p = 1\%$	$p = 10\%$	$p = 0,5\%$	$p = 1\%$	$p = 10\%$		
$m = 20$	$\mu = 3$	$\alpha = 0,1$	9,1580 (0,4012)	9,9113 (0,4296)	4,7895 (0,2568)	6,2061 (0,2446)	8,2697 (0,0934)	3,2966 (0,2261)	9,7957 (0,1291)	9,5909 (0,0371)	1,2409 (0,2930)	
		$\alpha = 0,01$	92,2430 (0,3894)	122,1134 (0,4631)	25,3668 (0,5264)	88,7185 (0,3748)	77,6925 (0,1355)	14,4394 (0,2233)	79,2859 (0,4907)	77,3266 (0,3468)	1,3568 (0,4758)	
	$\mu = 15$	$\alpha = 0,1$	9,9586 (0,2633)	8,9784 (0,3252)	2,0125 (0,3628)	10,0452 (0,1430)	2,4469 (0,2120)	8,9832 (0,5208)	8,1877 (0,1320)	8,6984 (0,2949)	1,3253 (0,5032)	
		$\alpha = 0,01$	96,2684 (0,1945)	94,2685 (0,4079)	20,4920 (0,4856)	87,6072 (0,1230)	108,7365 (0,2776)	5,9883 (0,2143)	90,0235 (0,4349)	89,3265 (0,4909)	1,9657 (0,0966)	
	$m = 50$	$\mu = 3$	$\alpha = 0,1$	10,1316 (0,3887)	10,1105 (0,3139)	5,6710 (0,1932)	10,1970 (0,1676)	10,1874 (0,2655)	4,3422 (0,3600)	9,9913 (0,1229)	7,8515 (0,3363)	1,3284 (0,4874)
			$\alpha = 0,01$	118,9603 (0,2735)	67,2478 (0,3662)	25,4840 (0,2211)	98,7405 (0,3280)	94,2359 (0,3194)	17,2699 (0,3491)	73,7635 (0,3246)	64,8411 (0,1037)	2,4897 (0,3373)
$\mu = 15$		$\alpha = 0,1$	8,2253 (0,2128)	8,5914 (0,4842)	3,2209 (0,3323)	10,0213 (0,4733)	9,8897 (0,5222)	6,4587 (0,5874)	9,5681 (0,2726)	8,3992 (0,1611)	1,3925 (0,2359)	
		$\alpha = 0,01$	76,9710 (0,1407)	48,2983 (0,4093)	14,6333 (0,4852)	99,7659 (0,3814)	96,2589 (0,3378)	15,3696 (0,3636)	91,0133 (0,1138)	89,2549 (0,2867)	3,0216 (0,2861)	
$m = 100$		$\mu = 3$	$\alpha = 0,1$	10,1352 (0,0818)	10,3804 (0,1564)	5,4031 (0,3091)	10,2686 (0,2814)	9,7790 (0,3142)	4,1989 (0,1849)	9,0532 (0,2032)	9,7059 (0,4276)	1,2340 (0,4835)
			$\alpha = 0,01$	93,8493 (0,2322)	123,0386 (0,1399)	18,4891 (0,1783)	99,9570 (0,2213)	95,4963 (0,1440)	13,7160 (0,1776)	88,4090 (0,3436)	59,9672 (0,2794)	2,4628 (0,0844)
	$\mu = 15$	$\alpha = 0,1$	9,6366 (0,2018)	9,8699 (0,3069)	4,3683 (0,4087)	9,0381 (0,2049)	10,1664 (0,3385)	2,5379 (0,1725)	9,3255 (0,1673)	9,7579 (0,4677)	1,1749 (0,2133)	
		$\alpha = 0,01$	81,2134 (0,2707)	89,0293 (0,0911)	14,6528 (0,0614)	77,2752 (0,2043)	107,0841 (0,1730)	10,5327 (0,3833)	80,5274 (0,1319)	100,1485 (0,0300)	1,6044 (0,3908)	

Tabela 10 – Valores médios de  $ARL_1$  (com desvios-padrões entre parênteses) do gráfico  $PSh_c$ , para os diferentes cenários estudados.

		$n = 50$			$n = 100$			$n = 500$				
		$p = 0,5\%$	$p = 1\%$	$p = 10\%$	$p = 0,5\%$	$p = 1\%$	$p = 10\%$	$p = 0,5\%$	$p = 1\%$	$p = 10\%$		
$m = 20$	$\mu = 3$	$\alpha = 0,1$	9,7986 (0,4318)	9,2785 (0,0882)	3,2357 (0,2462)	9,9453 (0,3955)	8,1769 (0,3940)	4,0059 (0,0302)	8,7152 (0,4396)	8,4484 (0,2308)	1,8401 (0,1873)	
		$\alpha = 0,01$	119,2171 (0,0834)	54,7354 (0,3234)	40,4403 (0,4633)	72,8640 (0,1157)	111,4875 (0,2260)	9,6657 (0,1303)	89,6357 (0,1914)	88,6276 (0,3638)	1,0155 (0,1498)	
	$\mu = 15$	$\alpha = 0,1$	9,9586 (0,1871)	8,9784 (0,3508)	2,0125 (0,1748)	10,0452 (0,4887)	7,7112 (0,2381)	2,4469 (0,1538)	8,1877 (0,1919)	8,6984 (0,1244)	1,3253 (0,4138)	
		$\alpha = 0,01$	96,2684 (0,0801)	94,2685 (0,2321)	20,4920 (0,1773)	87,6072 (0,3542)	108,7365 (0,2776)	5,9883 (0,1669)	90,0235 (0,2595)	89,3265 (0,3289)	1,9657 (0,3449)	
	$m = 50$	$\mu = 3$	$\alpha = 0,1$	9,6192 (0,4622)	9,3186 (0,1833)	4,5456 (0,4412)	9,0235 (0,3491)	8,6393 (0,4754)	3,0372 (0,1672)	9,9033 (0,4724)	9,6189 (0,4274)	1,3079 (0,3011)
			$\alpha = 0,01$	76,0216 (0,4613)	45,7224 (0,3309)	17,4058 (0,3031)	86,7895 (0,0849)	88,3659 (0,5046)	11,3270 (0,1205)	103,6386 (0,4971)	86,6013 (0,2157)	2,0058 (0,4762)
$\mu = 15$		$\alpha = 0,1$	9,7773 (0,1232)	9,3523 (0,1391)	3,4451 (0,4368)	9,6703 (0,1246)	10,1132 (0,4321)	5,6847 (0,6254)	9,4230 (0,1893)	8,8038 (0,2427)	1,0275 (0,3478)	
		$\alpha = 0,01$	51,5334 (0,5332)	113,6625 (0,4764)	14,1542 (0,3248)	88,7500 (0,3522)	87,3874 (0,0868)	11,3359 (0,2548)	103,4019 (0,1288)	82,0951 (0,2553)	2,2849 (0,3054)	
$m = 100$		$\mu = 3$	$\alpha = 0,1$	9,3977 (0,3502)	8,6123 (0,3687)	5,0512 (0,3459)	9,1930 (0,4021)	10,0222 (0,3361)	3,5966 (0,4731)	9,6993 (0,4219)	8,1562 (0,1868)	1,3320 (0,4897)
			$\alpha = 0,01$	96,5239 (0,4946)	65,6545 (0,3739)	24,2187 (0,3297)	83,8149 (0,5137)	77,7403 (0,1675)	14,1520 (0,4339)	87,2689 (0,0873)	12,1203 (0,4296)	2,6178 (0,2365)
	$\mu = 15$	$\alpha = 0,1$	8,7221 (0,4397)	9,1192 (0,3113)	8,8828 (0,4138)	10,2260 (0,5212)	10,1343 (0,2711)	2,5296 (0,3526)	8,5684 (0,4985)	9,1556 (0,3187)	1,1604 (0,3655)	
		$\alpha = 0,01$	62,9661 (0,3419)	68,6260 (0,2575)	10,8278 (0,5342)	97,1564 (0,1463)	91,2615 (0,2246)	8,2849 (0,4313)	71,0576 (0,2465)	66,5312 (0,1294)	1,5811 (0,5357)	

Tabela 11 – Valores médios de  $ARL_1$  (com desvios-padrões entre parênteses) do gráfico  $PSu_c$ , para os diferentes cenários estudados.

		$n = 50$			$n = 100$			$n = 500$				
		$p = 0,5\%$	$p = 1\%$	$p = 10\%$	$p = 0,5\%$	$p = 1\%$	$p = 10\%$	$p = 0,5\%$	$p = 1\%$	$p = 10\%$		
$m = 20$	$\mu = 3$	$\alpha = 0,1$	8,2454 (0,4663)	9,4759 (0,1218)	5,4787 (0,1929)	9,8025 (0,3201)	9,2502 (0,4779)	2,6213 (0,2951)	7,0659 (0,2123)	9,8595 (0,2856)	1,3478 (0,3909)	
		$\alpha = 0,01$	109,0121 (0,4755)	61,5839 (0,1244)	14,0874 (0,1833)	53,8984 (0,5271)	51,7190 (0,3031)	12,5751 (0,1433)	98,5122 (0,1387)	99,2226 (0,4377)	1,1845 (0,4296)	
	$\mu = 15$	$\alpha = 0,1$	9,3698 (0,4318)	9,9256 (0,4606)	3,6556 (0,3157)	9,6462 (0,1988)	5,7842 (0,3351)	3,1616 (0,1072)	9,2823 (0,3279)	8,4358 (0,1137)	2,4267 (0,3101)	
		$\alpha = 0,01$	81,7878 (0,1735)	66,9427 (0,2546)	7,0585 (0,2714)	94,1344 (0,2634)	73,2483 (0,5303)	5,0771 (0,5016)	66,3570 (0,1602)	68,2658 (0,1106)	2,2472 (0,1983)	
	$m = 50$	$\mu = 3$	$\alpha = 0,1$	10,2782 (0,3043)	10,0839 (0,3338)	4,8183 (0,4294)	10,1271 (0,1209)	10,0678 (0,3449)	3,7502 (0,1916)	9,7956 (0,1912)	9,5198 (0,2938)	9,3257 (0,1922)
			$\alpha = 0,01$	87,0433 (0,3231)	80,3273 (0,2585)	19,6956 (0,4970)	104,5672 (0,1040)	101,3660 (0,4590)	8,2369 (0,4463)	101,5119 (0,3486)	99,6879 (0,2400)	2,1898 (0,5194)
$\mu = 15$		$\alpha = 0,1$	9,7090 (0,2384)	9,9392 (0,0932)	3,4039 (0,3195)	9,3420 (0,2410)	10,0421 (0,1468)	7,4415 (0,7844)	9,2823 (0,2445)	8,4358 (0,3387)	1,3169 (0,4388)	
		$\alpha = 0,01$	88,1292 (0,2589)	107,2365 (0,8569)	13,4776 (0,1253)	101,5482 (0,1392)	99,3653 (0,1077)	11,2587 (0,2465)	51,2071 (0,1975)	49,2357 (0,1507)	2,2472 (0,0925)	
$m = 100$		$\mu = 3$	$\alpha = 0,1$	10,2373 (0,3895)	10,1320 (0,4121)	4,0820 (0,2265)	8,8571 (0,1337)	9,4110 (0,1221)	3,7579 (0,4221)	9,9977 (0,2419)	8,4093 (0,3226)	1,2764 (0,3637)
			$\alpha = 0,01$	86,9019 (0,1044)	103,4213 (0,3249)	23,7563 (0,3571)	82,7189 (0,1439)	102,2648 (0,3536)	14,1613 (0,3691)	99,6987 (0,1038)	94,8917 (0,3497)	2,4033 (0,3387)
	$\mu = 15$	$\alpha = 0,1$	9,7165 (0,1562)	9,9545 (0,3642)	3,4520 (0,4628)	9,9442 (0,3675)	9,3690 (0,2417)	2,2194 (0,2837)	9,2568 (0,1017)	9,3883 (0,3732)	1,0468 (0,3265)	
		$\alpha = 0,01$	89,0124 (0,6854)	98,2596 (0,4789)	12,0078 (0,5042)	81,2596 (0,9583)	88,1498 (0,5369)	7,7379 (0,1231)	80,2310 (0,7548)	99,6578 (0,5842)	1,3534 (0,4502)	

## 4 Aplicações

A teoria de controle de misturas uniparamétricas de Poisson introduzida aqui pode ser aplicada, por exemplo, em situações em que a natureza do processo de produção permite que um item ou produto apresente várias não-conformidades (ou defeitos), mas não seja classificado como não-conforme (ou defeituoso). Um bom exemplo dado por Montgomery (2013) é sobre a fabricação de computadores pessoais, que podem ter algumas pequenas falhas no acabamento do gabinete, mas serem rotulados como conformes, uma vez que essas falhas não afetam seriamente a operação funcional da unidade.

Os gráficos de controle propostos também podem ser vistos como alternativas atraentes e úteis para os conhecidos gráficos  $c$  e  $u$  derivados da suposição de Poisson, e os gráficos  $cm_{pc}$  e  $cm_{pu}$  desenvolvidos a partir da distribuição COM-Poisson, entre outros, no monitoramento de dados de contagem que exibem superdispersão.

Neste capítulo, são aplicadas as novas ferramentas CEP para inspecionar o número total ou médio de eventos por unidade de processo. Especificamente, são considerados conjuntos de dados artificiais (Seção 4.1), reais mais artificiais (Seção 4.2) e reais (Seção 4.3), contendo alguns níveis de superdispersão, para ilustrar a utilidade dos gráficos de controle usando as distribuições PL, PSh e PSu, bem como sua capacidade de gerar limites comparáveis aos estabelecidos pela teoria clássica de CEP.

Como no Capítulo 3, todas as análises foram realizadas utilizando o *software* R, e foi assumido  $L = 3$ .

### 4.1 Dados artificiais

Nesta aplicação, foram simuladas  $m = 50$  amostras de tamanhos desiguais (com  $\bar{n} = 100$ , e cada tamanho de amostra obtido seguindo a mesma abordagem descrita no primeiro parágrafo da Seção 3.1) de uma distribuição PL com parâmetro  $\mu = 3$ . A ideia aqui era comparar o desempenho do gráfico  $PL_u$  com alguns gráficos tradicionais do tipo  $u$ , especialmente na fase 2. De acordo com Sellers (2012), os gráficos de controle baseados na distribuição de Poisson (gráficos  $c$  e  $u$ ), apesar de apresentarem melhor performance para grandes amostras, não funcionam bem nos casos em que há superdispersão. Assim, a autora desenvolveu os gráficos de controle baseados na COM-Poisson (gráficos  $cm_{pc}$  e  $cm_{pu}$ ), usando esta distribuição de contagem mais geral que relaxa a suposição de equidispersão da distribuição de Poisson. Os gráficos de controle derivados da suposição de COM-Poisson têm mostrado excelentes resultados ao ajustar dados de contagem superdispersos. Portanto, os gráficos  $u$  e  $cm_{pu}$  podem ser usados como ferramentas CEP de referência para comparar o desempenho do gráfico  $PL_u$  nesta seção.

Na fase 1, são calculados os limites de controle dos três gráficos selecionados a partir das  $m = 50$  amostras mencionadas. Enquanto que na fase 2, são geradas  $m = 70$  novas amostras de tamanho variável ( $\bar{n} = 100$ , como antes) da mesma distribuição de PL, mas com as últimas 20 amostras perturbadas por  $\mu_s = 3,3$ . Assim, para todos os gráficos de controle construídos, foi possível detectar a ocorrência (ou não) de alarmes falsos nas primeiras 50 amostras, além do número de corridas até a primeira das 20 amostras não-conformes exceder os limites de controle.

A Figura 1 ilustra essa aplicação, em que os pontos (amostras) à direita da linha vertical tracejada verde são as observações perturbadas (amostras fora de controle). Além disso, as linhas horizontais sólidas vermelhas representam o LSC e o LIC, a linha horizontal sólida azul corresponde à LC e os pontos verdes indicam os sinais fora de controle.

A partir dessa figura, pode-se observar que, no caso de amostras conformes (sob controle), os gráficos  $PL_u$  (painel esquerdo) e  $cm_pu$  (painel direito) não produziram alarmes falsos, enquanto o gráfico  $u$  (painel do meio) detectou cinco alarmes falsos. No entanto, o gráfico de controle baseado na distribuição PL teve um desempenho ligeiramente superior ao do gráfico de controle baseado na distribuição COM-Poisson quando na presença de amostras não-conformes, em que as amostras PL fora de controle foram geradas com  $\mu_s = 3,3$ .

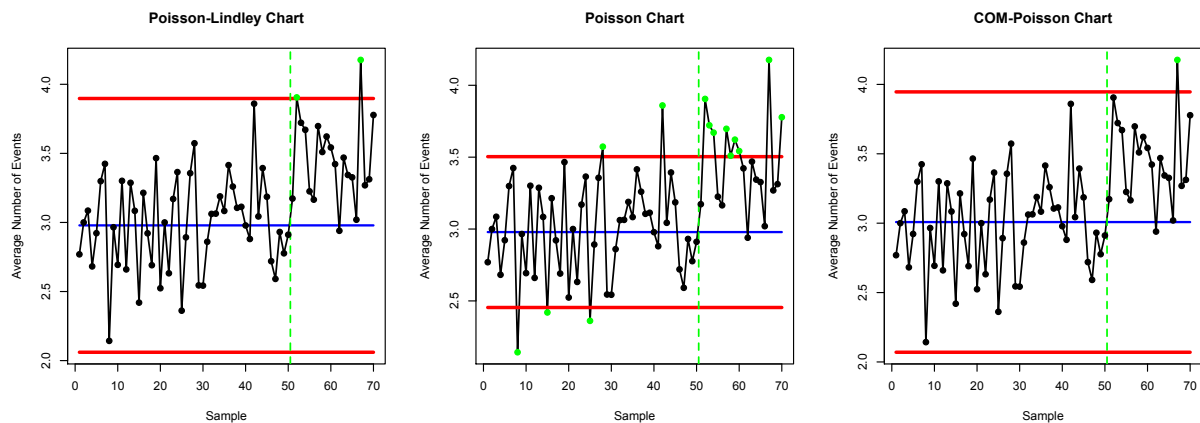


Figura 1 – Comparação de desempenho dos gráficos  $PL_u$ ,  $u$  e  $cm_pu$ , todos construídos a partir de amostras de distribuição PL (fase 2).

## 4.2 Dados reais mais artificiais

Nesta seção, são considerados os gráficos  $PSh_c$ ,  $c$  e  $BN-c$  para analisar o conjunto de dados reais relacionado ao número de erros de informação encontrados em registros de clientes na base de dados de uma empresa de marketing (ver Tabela 7E.23 de Montgomery, 2013). A análise dos dados abrange, inicialmente, apenas as últimas 15 das 25 amostras (dias), cada uma com 5 registros. Esses dados exibem superdispersão (variância/média

3,54), sugerindo que o processo pode ser descrito por uma distribuição PSh com  $\hat{\mu} = 6,01$ . Além disso, com base na abordagem de Alevizakos and Koukouvinos (2019), foram consideradas 15 amostras adicionais, com as primeiras 10 amostras simuladas de uma distribuição PSh com  $\mu = 6,01$  (amostras sob controle), e as últimas 5 amostras obtidas da mesma distribuição, mas com  $\mu_s = 6,61$  (amostras fora de controle). Isso implica em uma piora (ou deterioração) do processo, com sua taxa de não-conformidades aumentada em  $p = 10\%$ . Aqui, foi investigada a aderência da distribuição PSh ( $\hat{\mu} = 6,01$ ) por meio da aplicação do teste  $\chi^2$  de Pearson às 15 amostras de Montgomery de tamanho 5 cada, obtendo p-valor = 0,34.

A Figura 2 exibe os gráficos  $\text{PSh}_c$ ,  $c$  e  $\text{BN-}c$ , cujos limites de controle foram calculados a partir das 15 amostras iniciais. O lado direito da linha vertical tracejada laranja representa o monitoramento das 15 amostras restantes, que não participaram da construção dos limites de controle (fase 2), com a linha vertical tracejada verde separando as amostras sob controle (à esquerda) das amostras fora de controle (à direita).

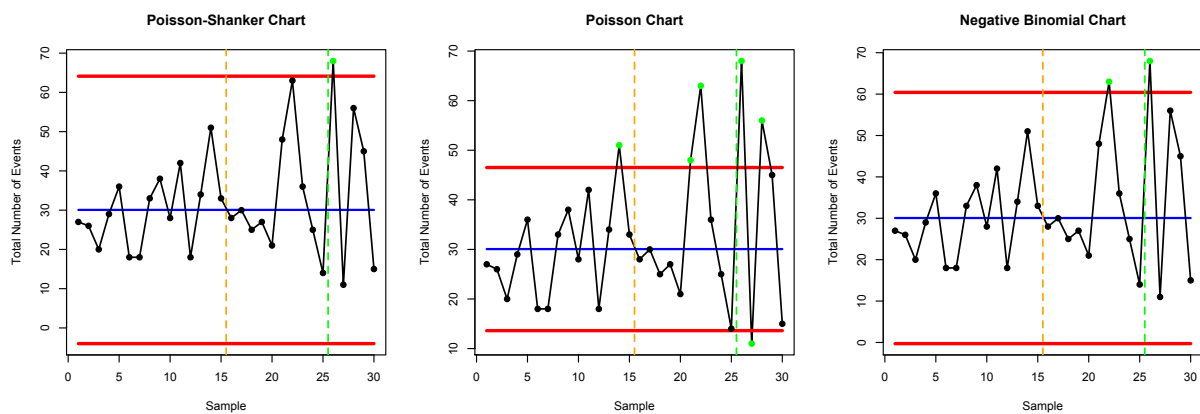


Figura 2 – Comparação de desempenho dos gráficos  $\text{PSh}_c$ ,  $c$  e  $\text{BN-}c$ , construídos a partir dos dados de erros de informação em registros de clientes.

Pode-se observar, a partir dessa figura, que os gráficos  $\text{PSh}_c$  e  $\text{BN-}c$  detectaram o deslocamento da média do processo para cima assim que ele ocorreu. Contudo, o gráfico  $\text{BN-}c$  produziu um único alarme falso na amostra (dia) #22, enquanto o gráfico  $\text{PSh}_c$  nenhum. Como esperado, o gráfico  $c$  gerou alarmes falsos para amostras de ambas as fases, 1 (amostra #14) e 2 (amostras #21 e #22). Isso ocorreu porque o gráfico  $c$  não é flexível o suficiente para lidar com dados de contagem superdispersos.

Portanto, os resultados obtidos sugerem que o gráfico  $\text{PSh}_c$  é uma excelente alternativa aos gráficos de controle tradicionais, como os gráficos  $c$  e  $\text{BN-}c$ .

### 4.3 Dados reais

Nesta seção, é ilustrada uma aplicação do gráfico de controle do tipo  $c$  com base na distribuição PSu, o chamado gráfico  $\text{PSu}_c$ , a dados reais sobre óbitos de COVID-19 (do inglês



*Coronavirus Disease 2019*), por semana epidemiológica (SE) de notificação, no estado da Bahia, Brasil. A ideia é ilustrar o uso das ferramentas CEP no monitoramento/vigilância epidemiológica de doenças infecciosas. Para outros trabalhos que usam ferramentas CEP para analisar dados da pandemia global de COVID-19, consulte, por exemplo, Perla et al. (2020), Singh et al. (2020) e Nnamdi et al. (2020).

A COVID-19 é uma doença respiratória que afeta as pessoas de maneiras diferentes. A maioria das pessoas infectadas com o vírus apresenta doença respiratória leve a moderada, recuperando-se sem a necessidade de tratamentos específicos (World Health Organization, 2020). Os principais sintomas da COVID-19 são febre, cansaço, tosse seca, falta de ar e dor de garganta. Idosos e aqueles com condições médicas subjacentes, como doenças cardiovasculares, diabetes, doenças respiratórias crônicas e câncer, têm maior probabilidade de desenvolver doenças graves e morrer. O vírus se espalha principalmente por meio de gotículas de saliva ou secreção nasal quando uma pessoa infectada tosse ou espirra. Segundo informações do consórcio de veículos de imprensa (formado por G1, “O Globo”, “Extra”, “O Estado de São Paulo”, “Folha de São Paulo” e UOL), até 19/05/2021, 19,06% da população brasileira (ou seja, 40.365.246 de pessoas) havia sido vacinada com a primeira dose de imunizantes, enquanto a segunda dose fora aplicada em 9,40% da população (19.909.403 de pessoas).

O conjunto de dados aqui analisado foi obtido do boletim epidemiológico disponível em [http://www.saude.ba.gov.br/wp-content/uploads/2017/11/BOLETIM-SRAG-14\\_2020.pdf](http://www.saude.ba.gov.br/wp-content/uploads/2017/11/BOLETIM-SRAG-14_2020.pdf). Compreende, inicialmente, apenas as SEs #19-#33 (fase 1). As contagens das SEs #34-#39 foram então usadas para fins de monitoramento (fase 2). Os dados das amostras da fase 1 revelam superdispersão (variância/média = 20,30), indicando que o processo pode ser descrito por uma distribuição PSu com  $\hat{\mu} = 348,33$  (p-valor = 0,6688 para o teste  $\chi^2$  de Pearson).

A Figura 3 mostra o gráfico PSu<sub>c</sub> gerado a partir dos dados de óbitos por COVID-19, em que a linha vertical tracejada laranja separa as amostras da fase 1 (análise retrospectiva) das amostras da fase 2 (análise prospectiva). Pode-se observar que o gráfico construído detectou um deslocamento médio para baixo (ou seja, uma melhoria de processo) da SE #18 (originalmente, SE #36) em diante. Além disso, esse gráfico mostrou um aumento acentuado nas mortes de COVID-19 da SE #8 até a SE #11 (originalmente, SEs #26-#29). Essas semanas compreendem as festas juninas (ou festas de São João), um período peculiar do Nordeste em que muitas pessoas da capital viajam para o interior. Após esse período, houve redução no número de óbitos por COVID-19.

Portanto, é imprescindível destacar que o gráfico PSu<sub>c</sub> proposto pode ser uma ferramenta CEP adequada para analisar e monitorar tais dados, indicando uma melhoria do processo (ou seja, uma redução significativa nas mortes por COVID-19) após um longo período estável.

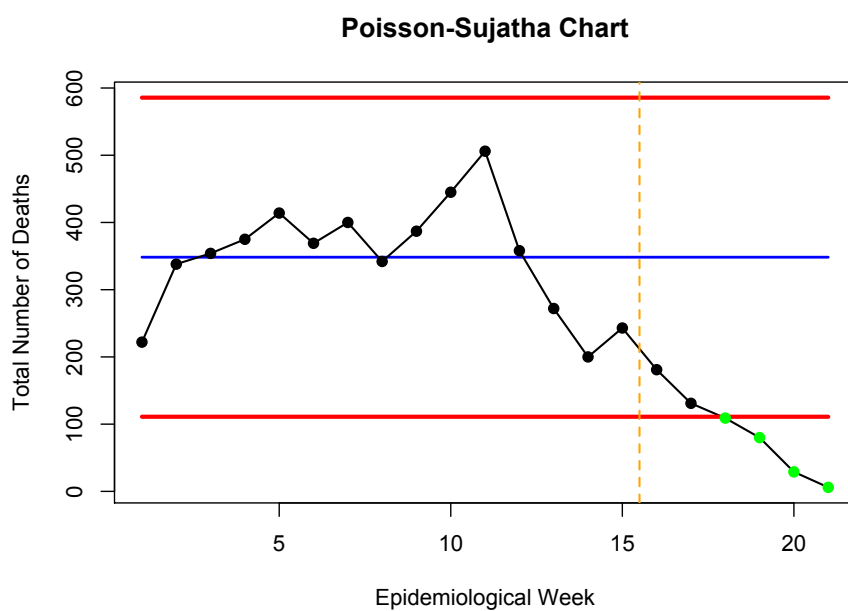


Figura 3 – Gráfico  $PSu_c$  para o número de óbitos de COVID-19, por semana epidemiológica de notificação, no estado da Bahia, Brasil.

## 5 Conclusões e Perspectivas Futuras

A classe de distribuições de misturas uniparamétricas de Poisson tem se mostrado uma alternativa viável à popular distribuição de Poisson para analisar dados de contagem com superdispersão. Conforme demonstrado pelos estudos empíricos e de simulação, os novos gráficos de controle baseados nessa classe de modelos discretos podem ser alternativas eficientes e úteis para alguns gráficos tradicionais (por exemplo, os gráficos  $c$  e  $u$  desenvolvidos a partir de distribuição de Poisson), ao lidar com dados de contagem superdispersos.

Existem algumas extensões possíveis do trabalho atual. Por exemplo, pode-se desenvolver um algoritmo de otimização a fim de melhorar o desempenho do gráfico de controle ao estimar o parâmetro de processo (desconhecido)  $\mu$  a partir dos dados da fase 1 (Braun, 1999; Chakraborti and Human, 2008; Castagliola and Wu, 2012; Castagliola et al., 2014). Além disso, a presença de covariáveis e autocorrelação não são incomuns na prática (Shu et al., 2014; Karaoglan and Bayhan, 2014). Levar em consideração essas características relevantes durante a modelagem nos permitiria, por exemplo, melhorar a aplicação do gráfico  $PSu_c$  aos dados sobre COVID-19, tornando esta ferramenta CEP também capaz de prever com precisão o futuro crescimento ou queda em casos ou óbitos pela doença, entre outros. Portanto, a abordagem apresentada deve ser investigada mais profundamente nesses contextos.

Por fim, vale destacar que um artigo resultante deste Trabalho de Conclusão de Curso, intitulado “*Statistical process control of overdispersed count data based on one-parameter Poisson mixture models*”, encontra-se atualmente resubmetido (após ter retornado com pedido de *Minor Revision*) para publicação em um importante periódico científico internacional.

# Referências

- Albers, W. (2011). Control charts for health care monitoring under overdispersion. *Metrica* 74(1), 67–83.
- Alevizakos, V. and C. Koukouvinos (2019). A progressive mean control chart for compoisson distribution. *Communications in Statistics-Simulation and Computation*, 1–19.
- Avcı, E., S. Altürk, and E. N. Soylu (2015). Comparison count regression models for overdispersed alga data. *International Journal of Research and Reviews in Applied Sciences* 25(1), 1–5.
- Boaventura, L. L., P. H. Ferreira, R. L. Fiaccone, P. L. Ramos, and F. Louzada (2020). New statistical process control charts for overdispersed count data based on the bell distribution. *Annals of the Brazilian Academy of Sciences (in press)*.
- Bourguignon, M., R. M. Medeiros, F. H. Fernandes, and L. Lee Ho (2020). Simple and useful statistical control charts for monitoring count data. *Quality and Reliability Engineering International*.
- Braun, W. J. (1999). Run length distributions for estimated attributes charts. *Metrica* 50(2), 121–129.
- Bulmer, M. G. (1974). On fitting the poisson lognormal distribution to species-abundance data. *Biometrics* 30(1), 101–110.
- Castagliola, P. and S. Wu (2012). Design of the c and np charts when the parameters are estimated. *International Journal of Reliability, Quality and Safety Engineering* 19(02), 1250010.
- Castagliola, P., S. Wu, M. B. C. Khoo, and S. Chakraborti (2014). Synthetic phase ii shewhart-type attributes control charts when process parameters are estimated. *Quality and Reliability Engineering International* 30(3), 315–335.
- Castellares, F., S. L. P. Ferrari, and A. J. Lemonte (2018). On the bell distribution and its associated regression model for count data. *Applied Mathematical Modelling* 56, 172–185.
- Chakraborti, S. and S. W. Human (2008). Properties and performance of the c-chart for attributes data. *Journal of Applied Statistics* 35(1), 89–100.
- Chen, J., L. Liu, L. Xiao, C. Xu, and D. Long (2020). Integrative analysis of spatial heterogeneity and overdispersion of crime with a geographically weighted negative binomial model. *ISPRS International Journal of Geo-Information* 9(1), 60.

- Cheng, S.-S. and F.-J. Yu (2013). A cusum control chart to monitor wafer quality. In *Proceedings of World Academy of Science, Engineering and Technology*, Number 78, pp. 1300. Citeseer.
- Coly, S., A.-F. Yao, D. Abrial, and M. Charras-Garrido (2016). Distributions to model overdispersed count data. *Journal de la Societe Francaise de Statistique* 157(2), 39–64.
- Conway, R. W. and W. L. Maxwell (1962). A queuing model with state dependent service rates. *Journal of Industrial Engineering* 12(2), 132–136.
- Famoye, F. (1994). Statistical control charts for shifted generalized poisson distribution. *Journal of the Italian Statistical Society* 3(3), 339–354.
- Ho, L. L., B. Andrade, M. Bourguignon, and F. H. Fernandes (2021). Monitoring count data with shewhart control charts based on the touchard model. *Quality and Reliability Engineering International*.
- Hougaard, P., M.-L. T. Lee, and G. A. Whitmore (1997). Analysis of overdispersed count data by mixtures of poisson variables and poisson processes. *Biometrics* 53(4), 1225–1238.
- Jensen, W. A., L. A. Jones-Farmer, C. W. Champ, and W. H. Woodall (2006). Effects of parameter estimation on control chart properties: a literature review. *Journal of Quality Technology* 38(4), 349–364.
- Kaminsky, F. C., J. C. Benneyan, R. D. Davis, and R. J. Burke (1992). Statistical control charts based on a geometric distribution. *Journal of Quality Technology* 24(2), 63–69.
- Karaoglan, A. D. and G. M. Bayhan (2014). A regression control chart for autocorrelated processes. *International Journal of Industrial and Systems Engineering* 16(2), 238–256.
- Karlis, D. and E. Xekalaki (2005). Mixed poisson distributions. *International Statistical Review/Revue Internationale de Statistique*, 35–58.
- Kokonendji, C. C. (2014). Over-and underdispersion models. In: *Methods and Applications of Statistics in Clinical Trials: Planning, Analysis, and Inferential Methods* (ed. N. Balakrishnan) 2, 506–526.
- Lindley, D. V. (1958). Fiducial distributions and bayes' theorem. *Journal of the Royal Statistical Society. Series B (Methodological)*, 102–107.
- McLachlan, G. J. and D. Peel (2004). *Finite mixture models*. John Wiley & Sons.
- Mohammed, M. A. and D. Laney (2006). Overdispersion in health care performance data: Laney's approach. *BMJ Quality & Safety* 15(5), 383–384.

- Montgomery, D. C. (2013). *Introduction to Statistical Quality Control* (7 ed.). John Wiley & Sons.
- Moraes, D., F. L. P. d. Oliveira, R. d. C. Quinino, and L. H. Duczmal (2014). Self-oriented control charts for efficient monitoring of mean vectors. *Computers & Industrial Engineering* 75, 102–115.
- Nnamdi, E., B. J. Odunayo, and O. D. Agbedeyi (2020). A retrospective analysis and monitoring of the spread of covid-19 in nigeria. *Global Journal of Pure and Applied Sciences* 26(2), 157–163.
- Nocedal, J. and S. Wright (2006). *Numerical Optimization*. Springer Science & Business Media.
- Okamura, H., A. E. Punt, and T. Amano (2012). A generalized model for overdispersed count data. *Population Ecology* 54(3), 467–474.
- Paroissin, C., L. Penalva, A. Pétrau, and G. Verdier (2016). New control chart for monitoring and classification of environmental data. *Environmetrics* 27(3), 182–193.
- Perla, R. J., S. M. Provost, G. J. Parry, K. Little, and L. P. Provost (2020). Understanding variation in covid-19 reported deaths with a novel shewhart chart application. *International Journal for Quality in Health Care*.
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. (Version 3.6.1). Vienna, Austria: R Foundation for Statistical Computing.
- R Core Team (2020). R: A language and environment for statistical computing, version 4.0.2. vienna, austria: R foundation for statistical computing. *Verfügbar unter <https://www.R-project.org>*.
- Saghir, A. and Z. Lin (2015). Control charts for dispersed count data: an overview. *Quality and Reliability Engineering International* 31(5), 725–739.
- Sankaran, M. (1970). The discrete poisson-lindley distribution. *Biometrics*, 145–149.
- Sellers, K. F. (2012). A generalized statistical control chart for over- or under-dispersed data. *Quality and Reliability Engineering International* 28(1), 59–65.
- Shaban, S. (1981, 01). On the discrete poisson-inverse gaussian distribution. *Biometrical Journal - BIOM J* 23, 297–303.
- Shanker, R. (2015). Shanker distribution and its applications. *International journal of statistics and Applications* 5(6), 338–348.

- Shanker, R. (2016a). The discrete poisson-amarendra distribution. *International Journal of Statistical Distributions and Applications* 2(2), 14.
- Shanker, R. (2016b, 01). On discrete poisson-shanker distribution and its applications. *Biometrics & Biostatistics International Journal* 5.
- Shanker, R. (2016c). Sujatha distribution and its applications. *Statistics in Transition. New Series* 17(3), 391–410.
- Shanker, R. (2017a). The discrete poisson-akash distribution. *International Journal of Probability and Statistics* 5(1), 1–9.
- Shanker, R. (2017b). The discrete poisson-aradhana distribution. *Türkiye Klinikleri Biyoistatistik* 9(1), 12–22.
- Shanker, R. (2017c). The discrete poisson-garima distribution. *Biometrics & Biostatistics International Journal* 5(2), 1–7.
- Shanker, R. and F. Hagos (2016). On poisson-sujatha distribution and its applications to model count data from biological sciences. *Biometrics & Biostatistics International Journal* 3(4), 1–7.
- Shanker, R., S. Sharma, R. Shanker, and T. Leonida (2014, 05). The discrete poisson-janardan distribution with applications. *International Journal of Soft Computing and Engineering* 4.
- Sheaffer, R. L. and R. S. Leavenworth (1976). The negative binomial model for counts in units of varying size. *Journal of Quality Technology* 8(3), 158–163.
- Shmueli, G., T. P. Minka, J. B. Kadane, S. Borle, and P. Boatwright (2005). A useful distribution for fitting discrete data: revival of the conway–maxwell–poisson distribution. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 54(1), 127–142.
- Shu, L., K.-L. Tsui, and F. Tsung (2014). Regression control charts. *Wiley StatsRef: Statistics Reference Online*.
- Shukla, K. and R. Shanker (2019). The discrete poisson-ishita distribution. *International Journal of Statistics and Economics* 20, 109–122.
- Singh, B. P., J. Madhusudan, A. K. Tiwari, S. Singh, and U. D. Das (2020). Evaluation of ewma control charts for monitoring spread of transformed observations of covid-19 in india. *Asian Journal of Research in Infectious Diseases*, 25–36.
- Spiegelhalter, D. J. (2005). Handling over-dispersion of performance indicators. *BMJ Quality & Safety* 14(5), 347–351.

World Health Organization (2020). *Coronavirus disease 2019: Pandemic*. World Health Organization.

Xekalaki, E. (2014). Under-and overdispersion. *Wiley StatsRef: Statistics Reference Online*, 1–9.

## Apêndice A

```
# samples1 = matrix with the total number of defects per phase 1 sample unit,
#           where each row represents a different sample
# L = "distance" between the center line and the control limits (in the usual
#     Six Sigma quality control program, L = 3)

PL.c.chart = function(samples1, L){
  y.BAR = apply(samples1, 1, mean)
  y.TOT = apply(samples1, 1, sum)
  n = ncol(samples1)
  mu = mean(y.BAR)
  a = -(1/(2*mu))*((mu-1)-sqrt(((mu-1)^2)+8*mu))
  desvio = sqrt(n*(a^3+(4*a^2)+6*a+2)/((a^2)*((a+1)^2)))
  UCL = n*mu+L*desvio
  CL = n*mu
  LCL = n*mu-L*desvio
  UCL = rep(UCL, length(y.TOT))
  CL = rep(CL, length(y.TOT))
  LCL = rep(LCL, length(y.TOT))
  plot(UCL, main="Poisson-Lindley c Chart", xlab="Sample", ylab="Total Number of Events",
       ylim=c(min(LCL, y.TOT), max(UCL, y.TOT)), col="red", type="l", lwd=3)
  lines(CL, col="blue", lwd=2)
  lines(y.TOT, col="black", lwd=1.5)
  lines(LCL, col="red", lwd=3)
  points(y.TOT, col = ifelse(y.TOT>UCL | y.TOT<LCL, "green", "black"), pch=19)
  out = matrix(round(c(UCL[1], CL[1], LCL[1]), 4), nrow=1, ncol=3)
  rownames(out) = c(" ")
  colnames(out) = c("UCL", "CL", "LCL")
  out
}

PSh.c.chart = function(samples1, L){
  y.BAR = apply(samples1, 1, mean)
  y.TOT = apply(samples1, 1, sum)
  n = ncol(samples1)
  mu = mean(y.BAR)
  s.mu = (12*mu*sqrt(3*(4*mu^4+71*mu^2+8))+180*mu^2+8)^(1/3)
```



```

a = (1/(3*mu))*(1+s.mu/2-2*(3*mu^2-1)/s.mu)
desvio = sqrt(n*((a^5+a^4+3*a^3+4*a^2+2*a+2)/(a^2*((a^2+1)^2)))
UCL = n*mu+L*desvio
CL = n*mu
LCL = n*mu-L*desvio
UCL = rep(UCL, length(y.TOT))
CL = rep(CL, length(y.TOT))
LCL = rep(LCL, length(y.TOT))
plot(UCL, main="Poisson-Shanker c Chart", xlab="Sample", ylab="Total Number of Events",
      ylim=c(min(LCL, y.TOT), max(UCL, y.TOT)), col="red", type="l", lwd=3)
lines(CL, col="blue", lwd=2)
lines(y.TOT, col="black", lwd=1.5)
lines(LCL, col="red", lwd=3)
points(y.TOT, col=ifelse(y.TOT>UCL | y.TOT<LCL, "green", "black"), pch=19)
out = matrix(round(c(UCL[1], CL[1], LCL[1]), 4), nrow=1, ncol=3)
rownames(out) = c(" ")
colnames(out) = c("UCL", "CL", "LCL")
out
}

```

```

h.mu = function(mu, log=F){
p1 = 3*sqrt(3)*sqrt(7*mu^4+28*mu^3+171*mu^2+32*mu+5)*mu
p2 = 8*mu^3+66*mu^2+6*mu+1
p = (p1+p2)^(1/3)
theta = log(p/mu-((mu-1)*(5*mu+1))/(mu*p)-(mu-1)/mu)-log(3)
if (log) theta else exp(theta)
}

```

```

PSu.c.chart = function(samples1, L){
y.BAR = apply(samples1, 1, mean)
y.TOT = apply(samples1, 1, sum)
n = ncol(samples1)
mu = mean(y.BAR)
a = h.mu(mu)
desvio = sqrt(n*(a^5+4*a^4+14*a^3+28*a^2+24*a+12)/(a^2*(a^2+a+2)^2))
UCL = n*mu+L*desvio
CL = n*mu
LCL = n*mu-L*desvio
UCL = rep(UCL, length(y.TOT))
CL = rep(CL, length(y.TOT))
LCL = rep(LCL, length(y.TOT))
plot(UCL, main="Poisson-Sujatha c Chart", xlab="Sample", ylab="Total Number of Events",
      ylim=c(min(LCL, y.TOT), max(UCL, y.TOT)), col="red", type="l", lwd=3)
lines(CL, col="blue", lwd=2)
lines(y.TOT, col="black", lwd=1.5)
}

```

```

lines(LCL, col="red", lwd=3)
points(y.TOT, col = ifelse(y.TOT>UCL | y.TOT<LCL, "green", "black"), pch=19)
out = matrix(round(c(UCL[1], CL[1], LCL[1]), 4), nrow=1, ncol=3)
rownames(out) = c(" ")
colnames(out) = c("UCL", "CL", "LCL")
out
}

```

```

PL.u.chart = function(samples1, L){
  y.TOT = apply(samples1, 1, sum, na.rm=TRUE)
  y.BAR = apply(samples1, 1, mean, na.rm=TRUE)
  counter = function(x){ sum(is.na(x)==FALSE) }
  N = apply(samples1, 1, counter)
  n = mean(N)
  mu = sum(y.TOT)/sum(N)
  a = -(1/(2*mu))*((mu-1)-sqrt(((mu-1)^2)+8*mu))
  desvio = sqrt((1/n)*(a^3+(4*a^2)+6*a+2)/((a^2)*((a+1)^2)))
  UCL = mu+L*desvio
  CL = mu
  LCL = mu-L*desvio
  UCL = rep(UCL, length(y.BAR))
  CL = rep(CL, length(y.BAR))
  LCL = rep(LCL, length(y.BAR))
  plot(UCL, main="Poisson-Lindley u Chart", xlab="Sample", ylab="Average Number of Events",
       ylim=c(min(LCL, y.BAR), max(UCL, y.BAR)), col="red", type="l", lwd=3)
  lines(CL, col="blue", lwd=2)
  lines(y.BAR, col="black", lwd=1.5)
  lines(LCL, col="red", lwd=3)
  points(y.BAR, col = ifelse(y.BAR>UCL | y.BAR<LCL, "green", "black"), pch=19)
  out = matrix(round(c(UCL[1], CL[1], LCL[1]), 4), nrow=1, ncol=3)
  rownames(out) = c(" ")
  colnames(out) = c("UCL", "CL", "LCL")
  out
}

```

```

PSh.u.chart = function(samples1, L){
  y.TOT = apply(samples1, 1, sum, na.rm=TRUE)
  y.BAR = apply(samples1, 1, mean, na.rm=TRUE)
  counter = function(x){ sum(is.na(x)==FALSE) }
  N = apply(samples1, 1, counter)
  n = mean(N)
  mu = sum(y.TOT)/sum(N)
  s.mu = (12*mu*sqrt(3*(4*mu^4+71*mu^2+8))+180*mu^2+8)^(1/3)
  a = (1/(3*mu))*(1+s.mu/2-2*(3*mu^2-1)/s.mu)
  desvio = sqrt((1/n)*((a^5+a^4+3*a^3+4*a^2+2*a+2)/(a^2*((a^2+1)^2))))

```

```

UCL = mu+L*desvio
CL = mu
LCL = mu-L*desvio
UCL = rep(UCL, length(y.BAR))
CL = rep(CL, length(y.BAR))
LCL = rep(LCL, length(y.BAR))
plot(UCL, main="Poisson-Shanker u Chart", xlab="Sample", ylab="Average Number of Events",
      ylim=c(min(LCL, y.BAR), max(UCL, y.BAR)), col="red", type="l", lwd=3)
lines(CL, col="blue", lwd=2)
lines(y.BAR, col="black", lwd=1.5)
lines(LCL, col="red", lwd=3)
points(y.BAR, col = ifelse(y.BAR>UCL | y.BAR<LCL, "green", "black"), pch=19)
out = matrix(round(c(UCL[1], CL[1], LCL[1]), 4), nrow=1, ncol=3)
rownames(out) = c(" ")
colnames(out) = c("UCL", "CL", "LCL")
out
}

```

```

PSu.u.chart = function(samples1, L){
  y.TOT = apply(samples1, 1, sum, na.rm=TRUE)
  y.BAR = apply(samples1, 1, mean, na.rm=TRUE)
  counter = function(x){ sum(is.na(x)==FALSE) }
  N = apply(samples1, 1, counter)
  n = mean(N)
  mu = sum(y.TOT)/sum(N)
  a = h.mu(mu)
  desvio = sqrt(((1/n)*(a^5+4*a^4+14*a^3+28*a^2+24*a+12))/(a^2*(a^2+a+2)^2))
  UCL = mu+L*desvio
  CL = mu
  LCL = mu-L*desvio
  UCL = rep(UCL, length(y.BAR))
  CL = rep(CL, length(y.BAR))
  LCL = rep(LCL, length(y.BAR))
  plot(UCL, main="Poisson-Sujatha u Chart", xlab="Sample", ylab="Average Number of Events",
        ylim=c(min(LCL, y.BAR), max(UCL, y.BAR)), col="red", type="l", lwd=3)
  lines(CL, col="blue", lwd=2)
  lines(y.BAR, col="black", lwd=1.5)
  lines(LCL, col="red", lwd=3)
  points(y.BAR, col = ifelse(y.BAR>UCL | y.BAR<LCL, "green", "black"), pch=19)
  out = matrix(round(c(UCL[1], CL[1], LCL[1]), 4), nrow=1, ncol=3)
  rownames(out) = c(" ")
  colnames(out) = c("UCL", "CL", "LCL")
  out
}

```

## Apêndice B

Tabela 12 – Valores médios de  $ARL_0$  (com desvio-padrões entre parênteses) dos gráficos  $PL_u$ ,  $PSh_u$  e  $PSu_u$ , para os diferentes cenários estudados,

		$n = 50$			$n = 100$			$n = 500$			
		$PL_c$	$PSh_c$	$PSu_c$	$PL_c$	$PSh_c$	$PSu_c$	$PL_c$	$PSh_c$	$PSu_c$	
$m = 20$	$\mu = 3$	$\alpha = 0, 1$	8,3569 (0,2960)	9,9321 (0,3373)	9,2589 (0,4899)	8,6029 (0,1283)	9,3661 (0,1295)	9,3682 (0,2620)	10,3929 (0,4068)	10,7180 (0,2735)	10,7046 (0,4574)
		$\alpha = 0, 01$	92,1059 (0,2952)	92,6462 (0,3162)	89,7902 (0,1264)	85,4676 (0,2475)	98,2214 (0,2185)	84,4503 (0,4489)	69,5874 (0,4255)	88,1215 (0,4864)	79,7659 (0,3328)
	$\mu = 15$	$\alpha = 0, 0027$	383,4455 (0,6182)	336,0607 (0,7460)	300,3489 (0,7361)	370,6541 (0,6153)	371,5682 (0,9025)	368,2547 (0,5355)	380,9685 (0,9899)	313,9584 (0,5105)	303,5812 (0,6103)
		$\alpha = 0, 1$	10,7589 (0,3142)	10,9653 (0,3484)	10,0244 (0,3321)	10,0873 (0,2530)	10,1483 (0,4888)	10,0334 (0,3663)	9,5655 (0,2987)	8,9976 (0,1032)	9,8568 (0,3994)
	$\mu = 3$	$\alpha = 0, 01$	105,7384 (0,3230)	99,1250 (0,2232)	99,1251 (0,1494)	102,1567 (0,2448)	97,7039 (0,2512)	98,1498 (0,3026)	89,2568 (0,2358)	99,1251 (0,4004)	77,2580 (0,1428)
		$\alpha = 0, 0027$	379,3160 (0,8419)	363,5262 (0,7111)	302,4115 (0,7249)	366,8957 (0,6415)	370,3254 (0,4381)	307,5689 (0,5386)	392,6941 (0,6875)	381,8754 (0,7723)	378,2378 (0,8924)
$m = 50$	$\mu = 3$	$\alpha = 0, 1$	9,8477 (0,1779)	9,2815 (0,3368)	9,6895 (0,4490)	9,3500 (0,0992)	9,2791 (0,5364)	9,6014 (0,5148)	9,5176 (0,3718)	9,5674 (0,2997)	9,4878 (0,4250)
		$\alpha = 0, 01$	99,1782 (0,1146)	90,9314 (0,3997)	95,9130 (0,3768)	94,4945 (0,2864)	82,3578 (0,3158)	88,5860 (0,2211)	78,1514 (0,3150)	83,1702 (0,1165)	75,7928 (0,5030)
	$\mu = 15$	$\alpha = 0, 0027$	393,0206 (0,6566)	372,6307 (0,9771)	376,5710 (0,8085)	398,1002 (0,4447)	394,4144 (0,8486)	378,0272 (0,6327)	381,0993 (0,8941)	373,5648 (0,8603)	368,1481 (0,7566)
		$\alpha = 0, 1$	10,1640 (0,1783)	9,3132 (0,4973)	9,5960 (0,2544)	9,1860 (0,1096)	9,7951 (0,1206)	9,1651 (0,2060)	9,5949 (0,3751)	9,5858 (0,4190)	9,6041 (0,3008)
	$\mu = 3$	$\alpha = 0, 01$	86,8460 (0,2014)	92,4929 (0,5164)	78,5143 (0,4938)	98,5369 (0,1746)	88,4069 (0,4968)	89,5856 (0,1552)	89,5422 (0,5016)	72,9803 (0,2038)	77,0431 (0,4114)
		$\alpha = 0, 0027$	396,3783 (0,9606)	393,0803 (0,9209)	373,9746 (0,6394)	399,8374 (0,4584)	385,3761 (0,8071)	390,6758 (0,7680)	399,3356 (0,7529)	363,5666 (0,9290)	313,2626 (0,5285)
$m = 100$	$\mu = 3$	$\alpha = 0, 1$	10,3720 (0,1310)	9,7768 (0,1114)	10,1849 (0,4603)	9,8332 (0,4231)	10,1315 (0,4677)	7,9347 (0,4126)	9,6291 (0,2809)	10,1139 (0,3048)	10,0256 (0,5187)
		$\alpha = 0, 01$	82,8135 (0,4080)	99,9645 (0,5170)	93,8423 (0,2189)	105,0058 (0,5343)	101,4245 (0,1521)	94,0516 (0,4472)	89,9034 (0,3373)	96,5155 (0,1042)	92,9522 (0,2986)
	$\mu = 15$	$\alpha = 0, 0027$	389,7481 (0,9648)	380,1931 (0,9220)	375,2761 (0,6236)	397,6005 (0,8118)	385,3780 (0,6072)	378,8651 (0,9764)	401,6638 (0,9777)	381,2072 (0,6929)	377,8200 (0,8666)
		$\alpha = 0, 1$	9,7419 (0,4018)	9,4943 (0,2759)	10,1932 (0,3742)	9,9262 (0,4854)	9,9697 (0,3752)	9,9240 (0,2407)	10,4588 (0,1447)	9,7226 (0,1515)	9,2357 (0,3791)
	$\mu = 3$	$\alpha = 0, 01$	98,5369 (0,0919)	88,4068 (0,5352)	89,5856 (0,2675)	95,6485 (0,2526)	98,1529 (0,2655)	95,0609 (0,3120)	95,2825 (0,2252)	91,6337 (0,1597)	88,0612 (0,2611)
		$\alpha = 0, 0027$	387,9416 (0,9518)	375,4006 (0,4091)	361,7770 (0,6565)	399,7291 (0,8527)	368,4848 (0,7253)	373,4947 (0,6772)	385,2991 (0,8714)	383,6434 (0,9134)	371,4367 (0,8986)

Tabela 13 – Valores médios de  $ARL_1$  (com desvios-padrões entre parênteses) do gráfico  $PL_u$ , para os diferentes cenários estudados,

		$\bar{n} = 50$			$\bar{n} = 100$			$\bar{n} = 500$				
		$p = 0,5\%$	$p = 1\%$	$p = 10\%$	$p = 0,5\%$	$p = 1\%$	$p = 10\%$	$p = 0,5\%$	$p = 1\%$	$p = 10\%$		
$m = 20$	$\mu = 3$	$\alpha = 0,1$	8,4473 (0,4919)	8,7584 (0,2417)	2,3652 (0,4853)	8,2750 (0,1348)	6,4134 (0,5689)	2,0168 (0,2205)	10,7512 (0,3228)	8,5892 (0,0848)	1,4072 (0,5339)	
		$\alpha = 0,01$	78,5632 (0,6845)	81,2547 (0,2658)	15,6681 (0,3063)	80,0065 (0,1511)	62,0787 (0,5181)	12,8882 (0,3279)	88,4680 (0,0227)	69,1658 (0,1810)	1,9530 (0,2086)	
	$\mu = 15$	$\alpha = 0,1$	9,9658 (0,3614)	9,5758 (0,1648)	2,4578 (0,0450)	9,8281 (0,0842)	9,5071 (0,4282)	2,4656 (0,5646)	9,5569 (0,2005)	9,1352 (0,0655)	1,2568 (0,2066)	
		$\alpha = 0,01$	96,9515 (0,3824)	56,8361 (0,2753)	19,4913 (0,4848)	89,7531 (0,1820)	94,6907 (0,3510)	15,6507 (0,0444)	58,3653 (0,1306)	57,6599 (0,4417)	4,0337 (0,4995)	
	$m = 50$	$\mu = 3$	$\alpha = 0,1$	9,8201 (0,3162)	9,7710 (0,3493)	5,6405 (0,0207)	9,5296 (0,1034)	9,6515 (0,5223)	4,6852 (0,0679)	9,4295 (0,1619)	10,0015 (0,5893)	1,4716 (0,3550)
			$\alpha = 0,01$	88,2514 (0,7789)	83,6512 (0,6754)	24,9368 (0,0384)	94,6250 (0,1049)	92,3569 (0,4278)	5,3660 (0,0580)	54,4570 (0,3806)	53,1441 (0,5636)	2,1347 (0,0710)
$\mu = 15$		$\alpha = 0,1$	9,9658 (0,5842)	9,5758 (0,0351)	2,4578 (0,4406)	8,3944 (0,2831)	10,0979 (0,3484)	2,9457 (0,2145)	8,5871 (0,3416)	8,1110 (0,4478)	1,3732 (0,3814)	
		$\alpha = 0,01$	71,6879 (0,2004)	63,5597 (0,2475)	12,5609 (0,1372)	99,6658 (0,5146)	97,4586 (0,3720)	10,3637 (0,4585)	52,8464 (0,3463)	58,3255 (0,1800)	3,0343 (0,4466)	
$m = 100$	$\mu = 3$	$\alpha = 0,1$	10,4045 (0,3386)	8,3770 (0,2208)	5,1087 (0,3187)	9,5706 (0,4931)	8,7173 (0,3252)	3,7284 (0,1378)	8,9614 (0,5119)	9,4813 (0,2349)	1,3831 (0,1959)	
		$\alpha = 0,01$	88,4569 (0,3569)	85,2789 (0,3698)	35,3274 (0,5066)	83,6531 (0,3500)	54,0847 (0,0352)	17,8279 (0,3002)	89,5578 (0,3568)	87,5632 (0,9865)	2,4648 (0,4426)	
	$\mu = 15$	$\alpha = 0,1$	9,4179 (0,3420)	9,1065 (0,0626)	3,4789 (0,0992)	2,7874 (0,4171)	9,8015 (0,0778)	2,5662 (0,1382)	9,4588 (0,2323)	9,7352 (0,5108)	1,1452 (0,1574)	
		$\alpha = 0,01$	88,2547 (0,9254)	80,7896 (0,6584)	12,6498 (0,4848)	76,2332 (0,1820)	71,6382 (0,3510)	10,2631 (0,0444)	80,1973 (0,7894)	88,2598 (0,5621)	1,6803 (0,4995)	

Tabela 14 – Valores médios de  $ARL_1$  (com desvios-padrões entre parênteses) do gráfico  $PSH_u$ , para os diferentes cenários estudados,

		$\bar{n} = 50$			$\bar{n} = 100$			$\bar{n} = 500$				
		$p = 0,5\%$	$p = 1\%$	$p = 10\%$	$p = 0,5\%$	$p = 1\%$	$p = 10\%$	$p = 0,5\%$	$p = 1\%$	$p = 10\%$		
$m = 20$	$\mu = 3$	$\alpha = 0,1$	9,3557 (0,5842)	9,6896 (0,0351)	7,8649 (0,4406)	9,2911 (0,2831)	9,8644 (0,3484)	6,7448 (0,5262)	10,6374 (0,3416)	8,7655 (0,4478)	1,5720 (0,3814)	
		$\alpha = 0,01$	93,2794 (0,3386)	55,3322 (0,2208)	51,4937 (0,3187)	97,5339 (0,4931)	84,9128 (0,3252)	10,3376 (0,1378)	88,2136 (0,5119)	64,9259 (0,2349)	2,1257 (0,1959)	
	$\mu = 15$	$\alpha = 0,1$	9,6955 (0,4669)	9,3528 (0,4454)	2,1458 (0,5066)	2,4656 (0,3500)	9,6365 (0,0352)	9,9849 (0,3020)	9,8046 (0,3348)	8,5655 (0,5053)	1,9956 (0,4426)	
		$\alpha = 0,01$	102,9442 (0,3420)	86,7446 (0,0626)	10,8960 (0,0992)	48,6191 (0,4171)	43,7593 (0,0778)	6,1660 (0,1382)	51,9366 (0,2323)	49,1077 (0,5108)	1,9856 (0,1574)	
	$m = 50$	$\mu = 3$	$\alpha = 0,1$	9,1204 (0,4288)	8,9070 (0,0824)	4,4177 (0,4719)	9,0250 (0,1680)	8,7230 (0,3379)	3,1980 (0,0397)	9,5243 (0,4875)	9,6924 (0,2159)	1,3677 (0,5514)
			$\alpha = 0,01$	106,1677 (0,2271)	81,2154 (0,5857)	17,1905 (0,2429)	69,7136 (0,3034)	69,3590 (0,3047)	3,6845 (0,0271)	85,2762 (0,3921)	44,9278 (0,3173)	3,1297 (0,2058)
$\mu = 15$		$\alpha = 0,1$	8,3269 (0,3572)	8,5844 (0,3781)	3,0930 (0,2742)	9,5033 (0,1188)	9,8996 (0,4347)	6,7893 (0,4978)	7,8442 (0,5095)	8,5349 (0,3302)	1,3302 (0,1953)	
		$\alpha = 0,01$	76,2528 (0,2848)	74,9783 (0,2429)	18,5769 (0,1644)	67,7256 (0,3997)	70,3366 (0,5741)	11,6258 (0,1888)	81,8935 (0,3971)	79,0755 (0,3311)	1,5184 (0,3665)	
$m = 100$	$\mu = 3$	$\alpha = 0,1$	9,6005 (0,3996)	8,1105 (0,4908)	4,8437 (0,2876)	10,1462 (0,1380)	9,6902 (0,0309)	4,0321 (0,1842)	9,5243 (0,5509)	9,6924 (0,5615)	1,3677 (0,3347)	
		$\alpha = 0,01$	92,8450 (0,1473)	61,3445 (0,3039)	23,7826 (0,4847)	98,1241 (0,2564)	89,1323 (0,0802)	16,1613 (0,1827)	88,2547 (0,4589)	70,5297 (0,9684)	2,6458 (0,2930)	
	$\mu = 15$	$\alpha = 0,1$	8,7801 (0,2322)	9,1393 (0,0903)	3,7748 (0,0473)	2,7210 (0,1724)	9,4715 (0,5818)	2,5114 (0,3033)	9,5914 (0,2971)	8,3823 (0,4543)	1,1479 (0,4072)	
		$\alpha = 0,01$	60,6589 (0,2145)	62,4513 (0,3578)	11,0010 (0,2307)	95,2396 (0,5328)	99,1453 (0,4697)	11,1080 (0,1007)	81,8935 (0,1911)	79,0755 (0,0931)	1,5184 (0,3621)	

Tabela 15 – Valores médios de  $ARL_1$  (com desvios-padrões entre parênteses) do gráfico  $PSu_u$ , para os diferentes cenários estudados,

		$\bar{n} = 50$			$\bar{n} = 100$			$\bar{n} = 500$				
		$p = 0,5\%$	$p = 1\%$	$p = 10\%$	$p = 0,5\%$	$p = 1\%$	$p = 10\%$	$p = 0,5\%$	$p = 1\%$	$p = 10\%$		
$m = 20$	$\mu = 3$	$\alpha = 0,1$	9,3570 (0,3458)	9,4735 (0,4195)	3,9586 (0,2005)	9,2701 (0,3712)	9,7436 (0,5948)	3,9891 (0,4511)	10,5051 (0,0640)	9,1790 (0,2817)	1,3251 (0,0511)	
		$\alpha = 0,01$	87,6697 (0,2169)	74,7492 (0,4457)	16,2295 (0,0224)	79,0026 (0,4677)	55,1987 (0,2885)	10,3562 (0,4381)	77,8215 (0,4066)	59,6863 (0,3518)	1,9766 (0,4282)	
	$\mu = 15$	$\alpha = 0,1$	9,3651 (0,4012)	9,7785 (0,1878)	1,7526 (0,0764)	9,2861 (0,5782)	9,7547 (0,4471)	1,8899 (0,3754)	8,8829 (0,0896)	8,4865 (0,3392)	1,6859 (0,1724)	
		$\alpha = 0,01$	102,9442 (0,5410)	86,7446 (0,0253)	8,9111 (0,1570)	85,7464 (0,0954)	90,5416 (0,2092)	8,6996 (0,4413)	59,6589 (0,5952)	69,2257 (0,4348)	2,6369 (0,3126)	
	$m = 50$	$\mu = 3$	$\alpha = 0,1$	9,7325 (0,2729)	9,7912 (0,5703)	5,6938 (0,0897)	9,5586 (0,0636)	9,3655 (0,5356)	3,4329 (0,2662)	9,2165 (0,0446)	8,5625 (0,3955)	1,2608 (0,2918)
			$\alpha = 0,01$	112,0270 (0,3784)	89,1310 (0,1771)	26,2533 (0,1112)	85,0630 (0,0863)	82,3298 (0,3145)	1,9541 (0,3379)	62,8979 (0,1016)	56,3863 (0,1185)	2,1906 (0,4620)
$\mu = 15$		$\alpha = 0,1$	8,8130 (0,3259)	9,6748 (0,5194)	3,9378 (0,4107)	8,1219 (0,0276)	10,1660 (0,4221)	6,1436 (0,3548)	7,9129 (0,3865)	9,4887 (0,0822)	1,1883 (0,5542)	
		$\alpha = 0,01$	108,1820 (0,4117)	72,7803 (0,1062)	12,0122 (0,4523)	89,0790 (0,5667)	88,3256 (0,2641)	8,2569 (0,1927)	41,9444 (0,1705)	72,2549 (0,1493)	2,5578 (0,3481)	
$m = 100$		$\mu = 3$	$\alpha = 0,1$	10,0146 (0,4589)	9,8988 (0,4084)	4,8461 (0,3370)	7,4618 (0,4906)	9,5481 (0,4603)	3,5854 (0,0316)	9,9993 (0,2409)	8,5390 (0,0495)	1,3509 (0,4828)
			$\alpha = 0,01$	89,2949 (0,5557)	55,7191 (0,3347)	21,1576 (0,5144)	95,5130 (0,3585)	81,7469 (0,4076)	18,6415 (0,3166)	87,7174 (0,4624)	45,6616 (0,5440)	2,4908 (0,4959)
	$\mu = 15$	$\alpha = 0,1$	9,8931 (0,0614)	10,0678 (0,0223)	3,6370 (0,0503)	9,8437 (0,5226)	8,9357 (0,1436)	2,2746 (0,2020)	9,4587 (0,5765)	9,3189 (0,3629)	1,0320 (0,3282)	
		$\alpha = 0,01$	88,2574 (0,2427)	95,0783 (0,2053)	12,7994 (0,4889)	57,3066 (0,0443)	94,9093 (0,2310)	5,9684 (0,5168)	54,0004 (0,4248)	71,1683 (0,4170)	1,2844 (0,2218)	