



Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Estatística

Dissertação de Mestrado

**Desenvolvimento e Aplicações do Processo
Estocástico de Touchard para
Contagens não-Poisson**

por

Moisés Lima

Brasília, 18 de Outubro de 2024

Desenvolvimento e Aplicações do Processo Estocástico de Touchard para Contagens não-Poisson

por

Moisés Lima

Dissertação apresentada ao Departamento de Estatística da Universidade de Brasília, como requisito parcial para obtenção do título de Mestre em Estatística.

Orientador: Prof. Dr. Raul Yukihiro Matsushita

Brasília, 18 de Outubro de 2024

Dissertação submetida ao Programa de Pós-Graduação em Estatística do Departamento de Estatística da Universidade de Brasília para a elaboração da dissertação de mestrado como um dos requisitos para a obtenção do título de Mestre em Estatística.

Texto aprovado por:

Prof. Dr. Raul Yukihiro Matsushita

Orientador, PPGEST/UnB

Profa. Dra. Regina Célia Bueno da Fonseca

PPGTGS/IFG

Prof. Dr. Gladston Luiz da Silva

PPGEST/UnB

Prof. Dr. Roberto Vila Gabriel

PPGEST/UnB

We have not got the money, so we have got to think.

(Ernest Rutherford)

Esta Dissertação é dedicada à minha avó Adia, conhecida por Sároza, e à minha esposa, Sifa.

À minha avó: agradeço por me teres educado da forma como o fizeste e por me teres encorajado sempre a levar os meus estudos o mais longe possível.

À minha esposa: Obrigado por acreditar sempre em mim e por me encorajar durante muitas longas noites de trabalho.

Agradecimentos

Este trabalho é o resultado de uma jornada repleta de desafios, aprendizagens e superações. A concretização deste projeto só foi possível graças ao apoio e incentivo de muitas pessoas, às quais gostaria de expressar minha profunda gratidão.

- Primeiramente, agradeço ao meu orientador, Prof. Dr. Raul Yukihiro Matsushita, por sua orientação incansável, paciência e conhecimento compartilhado ao longo deste percurso. Sua expertise e dedicação foram fundamentais para a realização deste trabalho.
- Aos meus professores do PPGEST/UnB, que proporcionaram ensinamentos essenciais tanto dentro quanto fora da sala de aula, expresso minha sincera gratidão.
- Expresso também minha gratidão aos meus colegas, amigos e ao grupo de pesquisa liderado pelo Prof. Dr. Rathie, que compartilharam momentos de estudo, discussões acadêmicas e apoio mútuo, tornando essa jornada mais leve e enriquecedora.
- Agradeço à minha família, por seu amor incondicional, apoio emocional e por sempre acreditarem em meu potencial. Sem vocês, nada disso seria possível.
- O presente trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Resumo

Apresentamos o processo estocástico de Touchard como um novo processo estocástico pontual pertencente à classe dos modelos de Poisson ponderados. Ele é obtido por meio da modificação das premissas do processo de Poisson tradicional, produzindo um sistema de equações de diferenciais, cuja solução está relacionada com os polinômios de Touchard. Entre as suas propriedades, o nosso modelo sugerido é um modelo flexível que pode ter em conta tanto a subdispersão ou superdispersão ou caudas pesadas e concentração de zeros que são frequentemente encontrados em dados não-Poisson. No entanto, ao contrário dos modelos estocásticos tradicionais, o processo de Touchard é gerado por incrementos não estacionários e dependentes, cujas probabilidades podem ser determinadas recursivamente. Além das suas propriedades, desenvolvemos um algoritmo de simulação para geração de contagens Touchard. Sua aplicabilidade é exemplificada através de dados de acidentes de trânsito nos condados de *Queens* e *Kings* do Estado de *New York* (EUA), sugerindo que o processo de Touchard seja um modelo candidato competitivo para o ajuste de contagens não-Poisson.

Palavras-chave: Processos de Poisson ponderados, superdispersão, processo de Poisson, processo de Touchard, subdispersão, distribuição inflacionada de zero

Abstract

We introduce the Touchard stochastic process as a new point stochastic process belonging to the class of weighted Poisson models. It is derived by changing the assumptions of the classical Poisson process, producing a system of difference differential equations whose solution is related to Touchard polynomials. Among its properties, our suggested model is flexible and can describe underdispersion, overdispersion, heavy tails, or concentration of zeros often found in non-Poisson data. However, unlike traditional stochastic models, the Touchard process is generated by non-stationary and dependent increments, whose probabilities can be obtained recursively. In addition to discussing its properties, we developed an algorithm for simulating Touchard counts. We illustrate its applicability with traffic accident data from *New York/USA* counties (*Queens* and *Kings*), showing that the Touchard process is a competitive candidate model for describing non-Poisson counts.

Keywords: Weighted Poisson processes, overdispersion, Poisson process, Touchard process, underdispersion, zero-inflated distribution

Sumário

1	Introdução	1
2	Distribuição de Poisson e suas Generalizações	4
2.1	Distribuição de Poisson	4
2.1.1	Distribuição de Poisson como Aproximação da Distribuição Binomial	7
2.2	Generalizações da Poisson	8
2.2.1	Poisson Generalizada	9
2.2.2	Distribuição Conway-Maxwell-Poisson	11
2.2.3	Modelo de Poisson Inflacionado de Zeros	13
2.2.4	Nova Poisson-Lindley Generalizada	15
2.2.5	Distribuição de Poisson Ponderada	16
2.2.6	Distribuição de Touchard	24
3	Processo de Poisson	29
3.1	Processo de Poisson Homogêneo	29
3.2	Função de Probabilidade para um Intervalo de Tempo de Comprimento	31
3.3	Equações de Diferença-Diferencial	31
4	Processo de Poisson Ponderado	36
4.1	Função de Peso Polinomial	39
4.1.1	$w(n) = n$	39

5	O Processo de Touchard	41
5.1	Derivadas Temporais	42
5.2	A taxa do PrT	46
5.3	Momentos	47
5.4	Fórmula recursiva para a função de probabilidade	49
5.5	Aspectos computacionais e simulação	50
6	Aplicação	53
6.1	Dados sobre acidentes com veículos motorizados	53
6.2	Observados no Condado de <i>Queens</i>	57
6.2.1	Parâmetros Estimados	61
6.2.2	Momentos Amostrais	61
6.2.3	Comparação de Modelos	61
6.2.4	A Estatística do Teste de Razão de Verossimilhança (TRV)	63
6.2.5	Significância dos Parâmetros ($\hat{\lambda}$ e $\hat{\delta}$)	63
6.3	Observados no Condado de <i>Kings</i>	64
7	Conclusão	70
A	Regressão de Poisson	72
A.1	Regressão de Poisson	72
A.2	Regressão Binomial	74
B	Outras Distribuições de Contagens	76
B.1	Binomial Negativa	76
B.2	Binomial Negativa Inflacionado com Zeros	78
	Referências Bibliográficas	79

Lista de Tabelas

2.1	Funções de peso frequentemente usadas	19
2.2	Funções de peso	19
2.3	Funções de peso não invertíveis	20
2.4	Funções de peso não invertíveis	21
6.1	Resumo das estimativas de máxima verossimilhança dos parâmetros do processo de Touchard e comparação com o modelo de Poisson no condado de <i>Queens</i>	60
6.2	Estimativas de máxima verossimilhança dos parâmetros do processo de Touchard e respectivos erros padrão (s.e.), a média amostral e a variância do número de acidentes observados no condado de <i>Kings</i> . Comparação dos Critérios de Informação Bayesiana (BIC) sob os modelos Touchard e Poisson, além do teste de Razão de Verossimilhança de Wilk (TRV) (H_0 : Poisson vs H_1 : Touchard), com os respectivos p -valores.	68
A.1	Parâmetros da distribuição Poisson	74
A.2	Parâmetros da distribuição Binomial	75

Lista de Figuras

5.1	Simulação de um processo de Touchard com $\lambda_t = 4t$, $\delta_t = 2$	52
6.1	Distribuição horário global (em horas).	55
6.2	Distribuição global por minuto.	55
6.3	Estimativas de máxima verossimilhança de λ_t , com $t = 5$ a 60 minutos respectivamente, com incremento de 5 minutos a partir das 18:00. As barras de erro referem-se aos intervalos de confiança de 99,7% com base na propriedade de normalidade assintótica.	56
6.4	Estimativas de máxima verossimilhança de δ_t , com $t = 5$ a 60, com incremento de 5 minutos a partir das 18:00. As barras de erro referem-se aos intervalos de confiança de 99,7% com base na propriedade de normalidade assintótica.	57
6.5	Variâncias do número de acidentes observados no condado de <i>Queens</i> (bolas sólidas), das 18:00 às 18:(00 + t), onde $t = 5$ a 60 minutos respectivamente, por incremento de 5 minutos. As bolas vazias denotam a média amostral correspondente, e a linha tracejada descreve sua evolução ao longo do tempo t , evidenciando a superdispersão em comparação com o processo de Poisson.	58
6.6	Barras verticais: Frequências relativas diárias de acidentes registados no condado de <i>Queens</i> , das 18:00 às 18:(00 + t), em que $t = 5, 10, 15, \dots, 60$ minutos respectivamente. As linhas pontilhadas destacam a forma da distribuição de Touchard (\circ).	59

6.7	Estimativas de máxima verossimilhança de λ_t , com $t = 5$ a 60 minutos respectivamente, com incremento de 5 minutos a partir das 17:00. As barras de erro referem-se aos intervalos de confiança de 99,7% com base na propriedade de normalidade assintótica.	64
6.8	Estimativas de máxima verossimilhança de δ_t , com $t = 5$ a 60, com incremento de 5 minutos a partir das 17:00. As barras de erro referem-se aos intervalos de confiança de 99,7% com base na propriedade de normalidade assintótica. . . .	65
6.9	Variâncias do número de acidentes observados no condado de <i>Kings</i> (bolas sólidas), das 17:00 às 18:(00 + t), onde $t = 5$ a 60 minutos respectivamente, por incremento de 5 minutos. As bolas vazias denotam a média amostral correspondente, e a linha tracejada descreve sua evolução ao longo do tempo t , evidenciando a superdispersão em comparação com o processo de Poisson.. . .	66
6.10	Barras verticais: Frequências relativas diárias de acidentes registados no condado de <i>Kings</i> , das 17:00 às 17:(00 + t), em que $t = 5, 10, 15, \dots, 60$ minutos respectivamente. As linhas pontilhadas destacam a forma da distribuição de Touchard (\circ).	67

Abreviações e Siglas

eq.	equação
Cov	Covariância
Var	Variância
TRV	teste de Razão de Verossimilhança
iid	Independentes e identicamente distribuídas
BIC	Crítérios de Informação Bayesiana
EMV	Estimadores de máxima verossimilhança
fdp	Função densidade de probabilidade
fmp	Função massa de probabilidade
FDA	Função de distribuição acumulada
FGM	Função geradora de momentos
DPG	Distribuição de Poisson generalizada
DCMP	Distribuição de Conway-Maxwell-Poisson
DPIZ	Distribuição de Poisson inflacionada com zeros
DPLG	Distribuição de Poisson de Lindley generalizada
DL	Distribuição de Lindley
DNPLG	Distribuição não condicional de Poisson-Lindley generalizada
DPP	Distribuição de Poisson ponderada

DT	Distribuição de Touchard
PrPP	Processo de Poisson ponderado
PrT	Processo de Touchard
MLG	Modelos lineares generalizados
FE	Família exponencial
DBNIZ	Distribuição binomial negativa inflacionado com zeros

Lista de Símbolos e Notações

N	Variável aleatória, Eq. (2.1)
$N(t)$	Variável aleatória de um processo
λ	Taxa de ocorrência, Eq. (2.1)
i, j, l, t	Índices
t	Janela de tempo em estudo
n	Tamanho de uma amostra
q	Ponto do suporte de uma FC
E	Esperança ou valor esperado
μ	Média
Var	Variância
σ^2	Variância
r	Rácio
$M(q)$	Momento de ordem q
$\dot{M}(q)$	Derivada do momento de ordem q
λ e δ	Parâmetros da distribuição de Touchard
$\tau(\lambda, \delta)$	Função normalizadora, Eq. (2.51)
$\dot{\tau}(\lambda_t, \delta_t)$	Derivada da função normalizadora
$p_n(t) = P[N(t) = n]$	Função de probabilidade para um intervalo de tempo de comprimento t Eq. (5.1)
$\text{Cov}_{s,t}$	Covariância do processo
$\dot{\mu}_t$	Função de taxa, Eq. (5.5)

$\dot{p}_n(t)$	Derivada temporal, Eq. (5.7)
$\pi_{n-1}(t)$	Probabilidade de transição de $n - 1$ para n do instante t para $t + \Delta t$, Eq. (5.8)
$\hat{\lambda}$	Estimativa de máxima verossimilhança do parâmetro λ do modelo de Touchard
$\hat{\delta}$	Estimativa de máxima verossimilhança do parâmetro δ_t do modelo de Touchard
s.e. $(\hat{\delta})$	Erro padrão da estimativa de δ
s.e. $(\hat{\lambda})$	Erro padrão da estimativa de λ
H_0 :	Hipótese nula
H_1 :	Hipótese alternativa
$\hat{\sigma}_t^2$	Variância da amostra para o número de acidentes observados
$\log \mathcal{L}$	O log da verossimilhança máxima

Capítulo 1

Introdução

O processo de Poisson é um modelo estocástico bastante conhecido que descreve contagens de ocorrências de um evento de interesse em tempo contínuo. Nesse modelo, esse evento é gerado espontaneamente, de forma aleatória, satisfazendo as suposições de independência e de estacionariedade dos incrementos. Como característica peculiar, as contagens de Poisson possuem média e variância coincidentes. Isto é, a sua utilização prática é limitada pela equidispersão o que significa que a média é igual à variância, fazendo com que a razão dessas quantidades, chamada de índice de Fisher ou índice de dispersão, seja igual a um. Mas, nem sempre que o índice de dispersão pode ser igual a um, pois, em algumas ocasiões, pode ser muito limitante. Os dados podem apresentar a superdispersão, ou seja, o índice de Fisher é maior que um. Para este caso, Já foram propostas distribuições para modelar os dados, como a distribuição binomial negativa, a distribuição ponderada de Poisson proposta por Del Castillo e Pérez-Casany (1998), o caso contrário é a subdispersão, onde o índice de Fisher é menor que um. A situação de subdispersão está bem documentada conforme exemplo dado por Daley e Maindonald (1989).

No entanto, há situações na prática nas quais o processo gerador das contagens apresentam estruturas de dependência temporal e não-estacionariedade que devem ser considerados na formulação do modelo. Isto é, há cenários do mundo real que muitas vezes exigem variações nas taxas de ocorrência dos eventos de interesse e na estrutura de dependências entre eles.

Por exemplo, se o tamanho de uma fila de espera por certo atendimento desencorajar a entrada de novos usuários nessa fila, então o tamanho da fila até determinado instante de tempo t poderia não seguir

uma distribuição de Poisson.

O processo de Poisson Ponderado pode servir como solução para situações, já que ele permite modelar a chegada de clientes com diferentes requisitos de tempo de serviço. Isso é fundamental para otimizar a alocação de recursos e a gestão de filas em empresas como centrais de atendimento, análise de tráfego de rede (Gross et al., 2011). Quando uma distribuição empírica apresenta uma superdispersão ou uma subdispersão — medida pelo índice de Fisher (empírico) ser relativamente maior ou menor do que um, respectivamente (Balakrishnan e Kozubowski, 2008) — dá-se a entender que nem sempre os incrementos são independentes ou estacionários. Por conta disso a literatura sugere outros modelos alternativos, que têm em conta tanto a superdispersão como a subdispersão. Balakrishnan e Kozubowski (2008) sugerem uma classe de processos estocásticos de Poisson ponderado gerada por uma taxa proporcional ao tempo e incrementos independentes associadas a uma função de peso.

No entanto, nem todos os processos estocásticos de Poisson Ponderado se enquadram nessa classe, entre eles, o processo de Touchard — que é o foco deste trabalho. Ele é um modelo flexível que permite acomodar contagens com superdispersão, subdispersão ou caudas pesadas que não pode ser gerada por incrementos estacionários e independentes. Em estatística e teoria das probabilidades, as distribuições discretas podem ser caracterizadas como tendo caudas "pesadas" ou "leves" com base na rapidez com que as probabilidades dos eventos extremos (valores afastados da média) diminuem. Uma distribuição discreta tem cauda pesada se a probabilidade dos valores extremos decresce mais lentamente. Isso significa que eventos extremos são mais prováveis do que seriam em uma distribuição de cauda leve (Embrechts, Klüppelberg e Mikosch, 2013).

Assim, o objetivo geral desta dissertação é estudar e desenvolver as propriedades do processo estocástico de Touchard para a descrição de contagens não-Poisson.

No que concerne à elaboração e organização desta dissertação, efetuamos uma revisão bibliográfica sobre a distribuição de Poisson e suas variações, incluindo a distribuição de Poisson ponderada e a de Touchard (Capítulo 2). Assuntos estudados complementarmente se encontram provisoriamente nos Apêndices A (Regressão de Poisson) e B (outras distribuições de contagens). Embora o processo de Poisson seja bastante conhecido, ele é apresentado no Capítulo 3, com ênfase na sua dedução por meio de uma equação denominada de equação de diferença-diferencial com base na premissa de incrementos

independentes e estacionários. O Capítulo 4 apresenta o processo de Poisson ponderado proposto por Balakrishnan e Kozubowski (2008) que se limita à classe dos modelos gerados por incrementos independentes e uma função de peso particular. No Capítulo 5 mostramos que o processo de Touchard não pertence à classe dos modelos propostos por Balakrishnan e Kozubowski (2008), e ainda mostramos como é possível obter recursivamente as probabilidades de transição das contagens. Mostramos que sob um intervalo de tempo $[0, t]$, um processo de Touchard evolui por meio de incrementos não-estacionários e dependentes, de modo que a contagem $N(t)$ se mantenha na forma de uma distribuição de Touchard com parâmetros indexados por t . Isso é feito com a ajuda do método da equação de diferença-diferencial do Capítulo 3, mas no sentido reverso. O Capítulo 6 apresenta a análise dos resultados. Por fim, no Capítulo 7 são apresentadas as conclusões, limitações do trabalho e recomendações para trabalhos futuros.

Capítulo 2

Distribuição de Poisson e suas Generalizações

Neste Capítulo, será feita uma revisão das principais distribuições de probabilidade utilizadas na modelagem de dados de contagem. Inicialmente, será apresentada a Distribuição de Poisson, amplamente utilizada em processos de contagem devido à sua simplicidade e aplicabilidade. Em seguida, será discutida a Distribuição Generalizada de Poisson, que oferece uma flexibilidade ao relaxar algumas das suposições da Poisson clássica. A Distribuição Conway-Maxwell Poisson também será abordada, uma alternativa que modela dados com superdispersão e subdispersão. A Distribuição de Poisson Inflacionada de Zeros será explorada como uma solução para lidar com dados de excesso de zeros, e, por fim, será apresentada a Distribuição de Touchard, uma extensão mais recente que permite uma modelagem ainda mais flexível de contagens que surgem de modificação de algumas suposições do modelo padrão Poisson. Essas distribuições fornecem uma base robusta para o desenvolvimento e aplicação de modelos em dados complexos de contagem.

2.1 Distribuição de Poisson

A Distribuição de Poisson desempenha um papel central na modelagem de dados de contagem. Trata-se de uma distribuição discreta que modela a probabilidade de ocorrência de um número fixo de eventos

em um intervalo de tempo ou espaço, considerando uma taxa média de ocorrência constante e eventos que ocorrem de forma independente. Proposta por Siméon Denis Poisson em 1838, essa distribuição tem sido amplamente utilizada em situações onde os eventos são raros e aleatórios (Ross, 2014).

Por ser a distribuição de Poisson o limite da distribuição binomial, ela tem sido amplamente utilizada na vida e nas ciências sociais, principalmente no ajuste da distribuição do número de eventos que só ocorrem com baixa probabilidade (Del Castillo e Pérez-Casany, 1998).

$$p(n) = P(N = n) = \frac{\lambda^n e^{-\lambda}}{n!}, \quad (2.1)$$

com $\lambda > 0$ e $n = 0, 1, 2, \dots$

Para verificar que a expressão (2.1) é uma distribuição de probabilidade legítima, temos que observar que

$$\sum_{n=0}^{\infty} P(N = n) = \sum_{n=0}^{\infty} \frac{\lambda^n e^{-\lambda}}{n!} = e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} = e^{-\lambda} e^{\lambda} = 1, \quad (2.2)$$

é uma série convergente, série de Maclaurin, ou seja, uma função exponencial.

Teorema 1. Se N tiver distribuição de Poisson com parâmetro $\lambda > 0$, então $E(N) = \lambda$ e $\text{Var}(N) = \lambda$, isto é: $E(N) = \text{Var}(N) = \lambda$ (e.g. Meyer, 1983).

Demonstração.

$$E(N) = \sum_{n=0}^{\infty} \frac{n\lambda^n e^{-\lambda}}{n!} = \sum_{n=1}^{\infty} \frac{n\lambda^n e^{-\lambda}}{n(n-1)!},$$

fazendo $s = n - 1$, verificamos que a expressão se torna

$$E(N) = \sum_{s=0}^{\infty} \frac{e^{-\lambda} \lambda^{s+1}}{s!} = \lambda \sum_{s=0}^{\infty} \frac{e^{-\lambda} \lambda^s}{s!} = \lambda. \quad (2.3)$$

Por definição, a variância é $\text{Var}(N) = E(N^2) - [E(N)]^2$. Assim,

$$E(N^2) = \sum_{n=0}^{\infty} \frac{n^2 \lambda^n e^{-\lambda}}{n!} = \sum_{n=0}^{\infty} \frac{n^2 \lambda^n e^{-\lambda}}{n(n-1)!} = \sum_{n=0}^{\infty} \frac{n \lambda^n e^{-\lambda}}{(n-1)!},$$

fazendo $s = n - 1$ e $s - 1 = t$ obtemos

$$\begin{aligned} E(N^2) &= \sum_{s=0}^{\infty} \frac{(s+1)\lambda^{s+1}e^{-\lambda}}{s!} = \lambda \sum_{s=0}^{\infty} \frac{s\lambda^s e^{-\lambda}}{s!} + \lambda \sum_{s=0}^{\infty} \frac{\lambda^s e^{-\lambda}}{s!} \\ &= \lambda \left[\sum_{s=1}^{\infty} \frac{\lambda^s e^{-\lambda}}{(s-1)!s} \right] + \lambda = \lambda e^{-\lambda} \lambda \sum_{t=0}^{\infty} \frac{\lambda^t}{t!} + \lambda = \lambda^2 e^{-\lambda} e^{\lambda} + \lambda = \lambda^2 + \lambda. \end{aligned} \quad (2.4)$$

Logo, pela definição

$$\text{Var}(N) = E(N^2) - [E(N)]^2 = [\lambda^2 + \lambda] - \lambda^2 = \lambda. \quad (2.5)$$

Podemos destacar outra característica importante das variáveis aleatórias discretas que é a Função Geradora de Momentos (FGM), definida para todos valores reais de t , e q variável discreta com função de probabilidade $p(k)$ (Magalhães, 2006).

$$M(q) = E[e^{qN}] = \sum_{k=0}^{\infty} e^{qk} p(k). \quad (2.6)$$

De acordo com Magalhães (2006), para $r = 1, 2, \dots$, o momento de ordem r da variável N é definido por $E(N^r)$, desde que essa quantidade exista. Se $E(N) = \mu < \infty$, definimos o momento central de ordem r por $E[(N - \mu)^r]$, sempre que essa quantidade exista. A equação (2.6) é denominado FGM, pois, todos os momentos de N podem ser obtidos com o cálculo sucessivo da derivada de $M(q)$ e com avaliação em $q = 0$, usando a equação (2.6).

$$M(q) = E[e^{qN}] = \sum_{k=0}^{\infty} e^{qk} p(k), \quad (2.6a)$$

$$E[q^r] = \frac{d^{(r)}}{dq^r} M(q),$$

quando $r = 1$, tem-se

$$\dot{M}(q) = \frac{d}{dq} E[e^{qN}] = E \left[\frac{d}{dq} e^{qN} \right] = E [N e^{qN}]. \quad (2.7)$$

De fato, da equação (2.6a), todos os momentos podem ser calculados com sucessivas derivadas de $M(q)$.

$$p(k) = \frac{e^{-\lambda} \lambda^k}{k!}$$

$$M(q) = E[e^{qN}] = \sum_{k=0}^{\infty} e^{qk} p(k) = E \left[\sum_{k=0}^{\infty} \frac{(qN)^k}{k!} \right] = \sum_{k=0}^{\infty} \frac{q^k}{k!} E[N^k]. \quad (2.8)$$

Para calcular a primeira derivada de $M(q)$ em relação a q , diferencie a série termo a termo e como a série é uma soma infinita de funções de q , a derivada da soma é a soma das derivadas. A derivada do termo geral $\frac{q^k}{k!} E[N^k]$ em relação a q é

$$\dot{M}(q) = \frac{d}{dq} \left(\sum_{k=0}^{\infty} \frac{q^k}{k!} E[N^k] \right) = \sum_{k=1}^{\infty} \frac{kq^{k-1}}{k!} E[N^k] = \sum_{k=1}^{\infty} \frac{q^{k-1}}{(k-1)!} E[N^k],$$

para simplificar ainda mais, fazemos a substituição $k' = k - 1$ temos

$$\dot{M}(q) = \sum_{k'=0}^{\infty} \frac{q^{k'}}{k'!} = \sum_{k'=0}^{\infty} \frac{q^{k'}}{k'!} E[N^{k'+1}] E[N^{k'}] = E \left[N \sum_{k'=0}^{\infty} \frac{(qN)^{k'}}{k'!} \right] = E [N e^{qN}],$$

portanto, a primeira derivada de $M(q)$ em relação a q é

$$\dot{M}(q) = E [N e^{qN}].$$

2.1.1 Distribuição de Poisson como Aproximação da Distribuição Binomial

O que acontece com as probabilidades binomiais se $m \rightarrow \infty$ e $p \rightarrow 0$, de tal maneira que np permaneça constante, quando $m \rightarrow \infty$ e $p \rightarrow 0$, de tal maneira que np permaneça constante? Tem-se

$$\begin{aligned} P(N = n) &= \binom{m}{n} p^n (1-p)^{m-n} \\ &= \frac{m!}{n!(m-n)!} p^n (1-p)^{m-n} \\ &= \frac{m(m-1)(m-2)\dots([m-(n+1)])}{n!} p^n (1-p)^{m-n}. \end{aligned} \quad (2.8a)$$

Para responder a questão acima deve-se saber que a distribuição de Poisson pode ser usada como uma aproximação para a distribuição binomial quando o número de tentativas m é grande e a probabilidade de sucesso p é pequena. Seja $mp = \lambda$, daí $p = \frac{\lambda}{m}$ e $1 - p = 1 - \frac{\lambda}{m}$, substituindo todos termos que contenha p por sua expressão equivalente em termos de λ , obtemos

$$\begin{aligned} P(N = n) &= \frac{m(m-1)(m-2)\dots(m-n+1)}{n!} \left(\frac{\lambda}{m}\right)^n \left(\frac{m-\lambda}{m}\right)^{m-n} \\ &= \frac{\lambda^n}{n!} \left[1 \left(1 - \frac{1}{m}\right) \left(1 - \frac{2}{m}\right) \dots \left(1 - \frac{n-1}{m}\right)\right] \left(1 - \frac{\lambda}{m}\right)^{m-n}. \end{aligned} \quad (2.8b)$$

Agora seja $m \rightarrow \infty$, de tal modo que $mp = \lambda$ permaneça constante enquanto $m \rightarrow \infty$, porque de outro modo mp não poderia ficar constante. Equivalentemente poderia-se impor que $m \rightarrow \infty$ e $p \rightarrow 0$, de tal modo que $mp \rightarrow \lambda$, na equação (2.8b)

$$\begin{aligned} \lim_{m \rightarrow \infty} P(N = n) &= \lim_{m \rightarrow \infty} \frac{m(m-1)(m-2)\dots(m-n+1)}{n!} \left(\frac{\lambda}{m}\right)^n \left(\frac{m-\lambda}{m}\right)^{m-n} \\ &= \frac{\lambda^n}{n!} \lim_{m \rightarrow \infty} \left[1 \left(1 - \frac{1}{m}\right) \left(1 - \frac{2}{m}\right) \dots \left(1 - \frac{n-1}{m}\right)\right] \left(1 - \frac{\lambda}{m}\right)^{m-n} = \frac{\lambda^n e^{-\lambda}}{n!}. \end{aligned}$$

Teorema 2. Se N uma variável aleatória distribuída binomialmente com parâmetro p (baseado em n repetições de um experimento) isto é: $P(N = n) = \binom{m}{n} p^n (1-p)^{m-n}$ admita-se que quando $m \rightarrow \infty$, fique $mp = \lambda$ (constante), ou equivalentemente, quando $m \rightarrow \infty$, $p \rightarrow 0$ de modo que $mp \rightarrow \lambda$. Nessas condições, tem-se: $\lim_{m \rightarrow \infty} P(N = n) = \frac{\lambda^n e^{-\lambda}}{n!}$ (e.g. Meyer, 1983).

2.2 Generalizações da Poisson

As generalizações da distribuição de Poisson referem-se a uma classe de distribuições que estendem o modelo básico da distribuição de Poisson para lidar com uma variedade mais ampla de cenários e padrões de contagem, para contornar os problemas de subdispersão, sobredispersão e excesso de zeros nos dados. Que anteriormente esses problemas acima referenciados foram detalhadamente tratados no que diz respeito ao seu conceito. Nesse capítulo vai-se abordar apenas de algumas das generalizações da distribuição de Poisson que foram desenvolvidas para atender às necessidades específicas de diferentes aplicações e contextos de modelagem. Essas generalizações são fundamentais na estatística e na mode-

lagem de dados de contagem, pois permitem que os analistas lidem com uma variedade mais ampla de cenários do mundo real.

2.2.1 Poisson Generalizada

De acordo com Lerner, Lone e Rao (1997), a distribuição de Poisson generalizada (DPG) é uma extensão da distribuição de Poisson clássica, que é usada para modelar a contagem de eventos raros em um intervalo fixo de tempo ou espaço. A importância da DPG reside no fato de que ela permite lidar com uma variedade mais ampla de cenários do que a distribuição de Poisson padrão. A distribuição de Poisson é uma das mais importantes distribuições de probabilidade. A DPG é uma distribuição de dois parâmetros definidas para $0 \leq \lambda \leq 1$, onde λ está relacionado com a dispersão e $\theta > 0$ controla a escala da distribuição, que é dada por

$$P(N = n) = \frac{\theta(\theta + n\lambda)^{n-1}e^{-n\lambda-\theta}}{n!}. \quad (2.9)$$

A média e a variância de N são dadas por

$$E(N) = \frac{\theta}{1 - \lambda}. \quad (2.10)$$

$$\text{Var}(N) = \frac{\theta}{(1 - \lambda)^3} = \frac{E(N)}{(1 - \lambda)^2}. \quad (2.11)$$

Interpretação da Relação entre θ e λ

A média e a variância da DPG, (2.10) e (2.11) respectivamente, tendem a aumentar ou diminuir conforme o valor do parâmetro λ aumenta ou diminui. Vejamos,

- À medida que θ aumenta, $E(N)$ aumenta linearmente.
- Quando λ se aproxima de 1, $E(N)$ aumenta, pois o denominador $1 - \lambda$ diminui. Isso indica que λ afeta a taxa de crescimento da média.
- À medida que θ aumenta, $\text{Var}(N)$ também aumenta.

- A variância cresce mais rapidamente do que a média à medida que λ aumenta, devido ao fator $(1 - \lambda)^3$ no denominador. Isso implica que λ tem um efeito mais pronunciado na variabilidade da distribuição.
- Quando λ está próximo de 0, a DPG se aproxima de uma distribuição de Poisson clássica, onde a média e a variância são aproximadamente iguais ($E(N) \approx \theta$ e $\text{Var}(N) \approx \theta$). Isto é, analisando a forma da função massa de probabilidade (fmp) da DPG e seu comportamento limite. Quando $\lambda \rightarrow 0$, o termo $(\theta + n\lambda)^{n-1} \approx \theta^{n-1}$, substituindo essa aproximação na fmp da DPG, tem-se

$$P(N = n) \approx \frac{\theta(\theta^{n-1})e^{-n\lambda-\theta}}{n!} = \frac{\theta^n e^{-n\lambda-\theta}}{n!}.$$

Agora quando $\lambda \rightarrow 0$, o termo $e^{-n\lambda}$ se aproxima de 1, portanto, a fmp da DPG se simplifica para

$$P(N = n) \approx \frac{\theta^n e^{-\theta}}{n!}.$$

A expressão $\frac{\theta^n e^{-\theta}}{n!}$ é exatamente a fmp de uma distribuição de Poisson com parâmetro θ . Na distribuição de Poisson clássica com parâmetro θ a média é $E(N) = \theta$ e a variância é $\text{Var}(N) = \theta$.

- À medida que λ aumenta, a média $E(N)$ e a variância $\text{Var}(N)$ aumentam, mas a variância aumenta mais rapidamente, indicando que a distribuição se torna mais dispersa.
- Para valores de λ próximos de 1, a variância pode se tornar muito grande, mesmo que a média não aumente proporcionalmente, destacando a flexibilidade da DPG em modelar dados com alta variabilidade.

Portanto, ao introduzir os parâmetros θ e λ , a DPG permite um ajuste mais preciso tanto da média quanto da variância dos dados modelados. Assim, a nova distribuição é tal que os dois parâmetros ajustam automaticamente os seus valores para se adequarem à maior parte dos padrões e fornecem aos investigadores um modelo que se pode ajustar facilmente aos seus dados, eliminando, de certa forma, uma das suas preocupações (Consul e Jain, 1973).

Os valores de θ e λ podem ser encontrados da seguinte forma

$$\theta = E(N)(1 - \lambda), \quad (2.12)$$

$$\lambda = 1 - \left(\frac{E(N)}{\text{Var}(N)} \right)^{\frac{1}{3}}. \quad (2.13)$$

2.2.2 Distribuição Conway-Maxwell-Poisson

A distribuição de Poisson é uma das distribuições discretas mais utilizadas, uma vez que os dados em campos de investigação cumprem frequentemente os postulados de Poisson. No entanto, a sua dependência de um único parâmetro limita a sua flexibilidade em muitas aplicações. A superdispersão dos dados relativamente à distribuição de Poisson é um problema frequente (Shmueli et al., 2005). A distribuição de Conway-Maxwell-Poisson (DCMP) que foi proposta por Conway e Maxwell (1962) e depois foram introduzidas e exploradas as suas propriedades estatísticas e probabilísticas por Shmueli et al. (2005), que caracterizou a distribuição de um ponto de vista probabilística e estatística (Boatwright et al., 2006). Portanto, trata-se de uma distribuição flexível que pode ter em conta a superdispersão ou a subdispersão que na vida real é comum encontrar-se em dados de contagem.

Função de probabilidade de Conway-Maxwell-Poisson

A DCMP generaliza a distribuição de Poisson, permitindo a superdispersão ou subdispersão. A sua função de probabilidade é dada por

$$P(N = n) = \frac{\lambda^n}{(n!)^\nu} \frac{1}{Z(\lambda, \nu)}, \quad (2.14)$$

$$n = 0, 1, 2, \dots,$$

$$Z(\lambda, \nu) = \sum_{j=0}^{\infty} \frac{\lambda^j}{(j!)^\nu}, \quad (2.15)$$

para $\lambda > 0$ e $\nu \geq 0$ (mas veja a exceção abaixo). Isto satisfaz as condições para uma função de

probabilidade.

Proposição 1. *A formulação (2.16) permite uma diminuição não linear dos rácios das probabilidades sucessivas na forma*

$$\frac{P(N = n - 1)}{P(N = n)} = \frac{n^\nu}{\lambda}. \quad (2.16)$$

Demonstração.

$$\frac{P(N = n - 1)}{P(N = n)} = \frac{\lambda^{n-1}(n!)^\nu Z(\lambda, \nu)}{[(n - 1)!]^\nu Z(\lambda, \nu)\lambda^n} = \lambda^{-1}n^\nu = \frac{n^\nu}{\lambda}. \quad (2.17)$$

□

Na equação (2.17), verifica-se que a série $\frac{\lambda^j}{(j!)^\nu}$ convergem para qualquer $\lambda > 0$ e $\nu > 0$, uma vez que o rácio dos dois termos subsequentes da série $\frac{\lambda}{(j)^\nu}$ tende para zero à medida que $j \rightarrow \infty$. A DCMP é uma generalização de algumas distribuições discretas bem conhecidas. No caso de $\nu = 1$, portanto $Z(\lambda, \nu) = e^\lambda$ resulta uma distribuição de Poisson(λ). Como $\nu \rightarrow \infty$, $Z(\lambda, \nu) \rightarrow 1 + \lambda$, e a DCMP aproxima-se de uma distribuição de Bernoulli com $P(N = 1) = \frac{\lambda}{1+\lambda}$. Quando $\nu = 0$ e $\lambda < 1$, $Z(\lambda, \nu)$ é uma soma geométrica,

$$Z(\lambda, \nu) = \frac{1}{1 - \lambda}, \quad (2.18)$$

e a própria distribuição é geométrica:

$$P(N = n) = \lambda^n(1 - \lambda), \quad (2.19)$$

para $n = 0, 1, 2, \dots$.

Os momentos da DCMP pertence à família das distribuições em série de potências de dois parâmetros (Shmueli et al., 2005). Os momentos desta distribuição podem então ser expressos utilizando a fórmula recursiva

$$E(N^{r+1}) = \begin{cases} \lambda E(N+1)^{1-\nu}, & \text{se } r = 0, \\ \lambda \frac{d}{d\lambda} E(N^r) + E(N) E(N^r) & \text{se } r > 0. \end{cases}$$

2.2.3 Modelo de Poisson Inflacionado de Zeros

A concentração de zeros são frequentemente encontrados em dados de contagem, os dados de contagens não-Poisson abrangem também os casos de Inflação de Zero. Os modelos inflacionados com zero são usados quando há mais zeros nos dados do que o esperado. Lambert (1992) introduziu conceito de inflação de zero para resolver o problema de excesso de zeros.

Poisson Inflacionado com Zeros

A distribuição de Poisson inflacionados com zero (DPIZ) combinam dois componentes, um para contagem de zero para explicar a probabilidade de uma observação ser igual à zero com probabilidade p e no segundo componente a própria distribuição de Poisson(λ) com probabilidade $1 - p$ (Hall, 2000). De acordo com Lambert (1992) e Viviano, Muggero e Lovison (2005), a sua expressão é dada por

$$f(n, p, \mu) = \begin{cases} p + (1 - p)e^{-\mu}, & \text{se } n = 0, \\ (1 - p) \frac{e^{-\mu} \mu^n}{n!} & \text{se } n > 0. \end{cases}$$

Na qual $\mu = \lambda$ é a média da distribuição de Poisson e p representa a probabilidade da ocorrência de zeros. A média da DPIZ e a sua variância podem ser encontradas de seguinte forma, (Lambert, 1992).

$$\begin{aligned} E(N) &= \sum_{n=0}^{\infty} n f(n, p, \mu) \\ &= \sum_{n=0}^{\infty} n (1 - p) \frac{e^{-\mu} \mu^n}{n!} \\ &= (1 - p) e^{-\mu} \sum_{n=0}^{\infty} n \frac{\mu^n}{n!} \\ &= (1 - p) e^{-\mu} \sum_{n=1}^{\infty} \frac{\mu^n}{(n-1)!}, \end{aligned}$$

fazendo $s = n - 1$, verifica-se que a expressão se torna

$$E(N) = (1 - p) e^{-\mu} \sum_{s=0}^{\infty} \frac{\mu^{s+1}}{s!} = (1 - p) \mu. \quad (2.20)$$

Para a variância, tem-se que encontrar em primeiro passo $E(N^2)$

$$\begin{aligned} E(N^2) &= \sum_{n=0}^{\infty} n^2 f(n, p, \mu) \\ &= \sum_{n=0}^{\infty} n^2 (1-p) \frac{e^{-\mu} \mu^n}{n!} \\ &= (1-p) e^{-\mu} \sum_{n=0}^{\infty} n^2 \frac{\mu^n}{n!} \\ &= (1-p) e^{-\mu} \sum_{n=1}^{\infty} \frac{n \mu^n}{(n-1)!}, \end{aligned}$$

fazendo $s = n - 1$, obtemos

$$\begin{aligned} E(N^2) &= (1-p) e^{-\mu} \sum_{s=0}^{\infty} \frac{(s+1) \mu^{s+1}}{s!} \\ &= \mu (1-p) e^{-\mu} \sum_{s=0}^{\infty} \frac{s \mu^s}{s!} + \mu (1-p) e^{-\mu} \sum_{s=0}^{\infty} \frac{\mu^s}{s!} \\ &= \mu (1-p) e^{-\mu} \sum_{s=1}^{\infty} \frac{s \mu^s}{s(s-1)!} + \mu (1-p), \end{aligned}$$

fazendo agora $s - 1 = t$, temos

$$\begin{aligned} E(N^2) &= \mu (1-p) e^{-\mu} \sum_{t=0}^{\infty} \frac{\mu^{t+1}}{t!} + \mu (1-p) \\ &= \mu^2 (1-p) + \mu (1-p) = (1-p)(\mu^2 + \mu). \end{aligned} \tag{2.21}$$

Portanto, a variância, $\text{Var}(N)$ é dada por

$$\begin{aligned} \text{Var}(N) &= E(N^2) - [E(N)]^2 = (1-p)(\mu^2 + \mu) - [(1-p)\mu]^2 \\ &= (\mu^2 + \mu - p\mu^2 - p\mu) - (\mu - p\mu)^2 \\ &= (\mu^2 + \mu - p\mu^2 - p\mu) - (\mu^2 - 2p\mu^2 - p^2\mu^2) \\ &= \mu(1-p)(1 + p\mu). \end{aligned} \tag{2.22}$$

2.2.4 Nova Poisson-Lindley Generalizada

A distribuição de Poisson-Lindley Generalizada (DPLG) surge da mistura de duas distribuições (Poisson e distribuição de Lindley) e foi proposta como uma alternativa flexível para modelar dados de contagem em que há superdispersão (variabilidade maior do que a esperada pela distribuição de Poisson) (Bhati, Sastry e Qadri, 2015). Shanker, Sharma e Shanker (2013) propuseram novas classes de distribuições modificando a distribuição de Lindley e discutiram várias propriedades das generalizações propostas. É considerando que o parâmetro λ da distribuição de Poisson segue a distribuição Lindley com dois parâmetros ($DL(\theta, \alpha)$), que foi sugerida por Shanker, Sharma e Shanker (2013), a sua função de densidade de probabilidade para a DPLG pode ser expressa por

$$f(n, \alpha, \theta) = \frac{\theta^2}{(\theta + \alpha)}(1 + \alpha n)e^{-\theta n}, \quad (2.23)$$

com $n, \alpha, \theta > 0$.

De acordo com Bhati, Sastry e Qadri (2015), diz-se que uma variável aleatória N segue a nova distribuição de Poisson-Lindley generalizada se ela seguir a representação estocástica

$$\begin{aligned} N|\lambda &\sim \text{Poisson}(\lambda), \\ \lambda|\theta, \alpha &\sim \text{DL}(\theta, \alpha), \end{aligned}$$

para $\lambda > 0$ e $\theta, \alpha > 0$. Denotamos a distribuição não condicional por $DNPLG(\theta, \alpha)$.

Teorema 3. Se $N \sim DNPLG(\theta, \alpha)$, então a função de massa de probabilidade de N é:

$$g(n, \alpha, \theta) = \frac{\theta^2}{(\theta + \alpha)(1 + \theta)^{n+1}} \left(1 + \frac{\alpha(n + 1)}{1 + \theta} \right), \quad (2.24)$$

$n = 1, 2, \dots$ e $\alpha, \theta > 0$.

Método dos momentos

As estimativas dos momentos, $\hat{\alpha}$ e $\hat{\theta}$, de α e θ podem ser obtidas resolvendo as seguintes equações

dos dois primeiros momentos

$$m_1 = \mu'_1 = \frac{2\alpha + \theta}{\theta(\alpha + \theta)}, \quad (2.25)$$

$$m_2 = \mu'_2 = \frac{2\alpha(\theta + 3) + \theta(\theta + 2)}{\theta^2(\alpha + \theta)}. \quad (2.26)$$

Resolvendo ambas equações, obtemos

$$\hat{\alpha} = \frac{\hat{\theta} - m_1 \hat{\theta}^2}{m_1 \hat{\theta} - 2}, \quad (2.27)$$

$$\hat{\theta} = \frac{2m_1 + \sqrt{4\mu_1^2 + 2m_1 - 2m_2}}{m_2 - m_1}. \quad (2.28)$$

2.2.5 Distribuição de Poisson Ponderada

O conceito de distribuição ponderada foi introduzida por Snyder, Neugebauer e Rao (1963). A distribuição de Poisson ponderado ocorre naturalmente em contextos onde a probabilidade de que uma observação particular da variável de Poisson entre na amostra é multiplicada por alguma função de peso não negativa (Balakrishnan e Kozubowski, 2008). A distribuição de Poisson ponderada (DPP) é uma alternativa para a distribuição de Poisson quando há subdispersão ou superdispersão (Patil e Rao, 1978). Isso nos permite ter em conta, entre outras situações da dispersão (Mizere et al., 2013), (Mizère, Kokonendji e Dossou-Gbété, 2006). O procedimento geral para obter a DPP é multiplicar a distribuição de Poisson pela razão da função peso pela constante de normalização (Kokonendji, Mizere e Balakrishnan, 2008), (Balakrishnan e Kozubowski, 2008). Isto é,

$$p(n)_w = P(N_w = n) = \frac{w(n)p(n)}{E[w(N)]}, \quad (2.29)$$

na qual $w(n)$ é uma função, $p(n)$ é a probabilidade de n , e $E[w(N)]$ é o valor esperado de $w(N)$, em que $n = 0, 1, 2, \dots$ e $w(n)$ é uma função de peso e não negativa, $p(n)$ é a função massa de probabilidade de Poisson e, $0 < E[w(N)] < \infty$ é a chamada constante de normalização. A função de peso pode depender de um parâmetro que representa o mecanismo de gravação, e pode também estar conectado ao parâmetro de Poisson subjacente (Balakrishnan e Kozubowski, 2008). A função de peso mais comum é

$w(n) = n, \forall n \in \mathbb{N}$, onde ela não depende do parâmetro de Poisson (Patil e Rao, 1978).

A DPP pode ser usada para modelar dados de contagem, que pode ser usada para modelar o número de vezes que um evento ocorre em um determinado período de tempo. Por exemplo, ela pode ser usada para modelar o número de acidentes de carro em um determinado trecho de estrada em um determinado período de tempo (Smith2020). Pode ser aplicada para dados de tempo de espera, por exemplo, ela pode ser usada para modelar o tempo que leva para um cliente ser atendido em um restaurante. Também em dados de incidência, por exemplo, ela pode ser usada para modelar o número de casos de COVID-19 em um país em um determinado período de tempo. É conhecida pela sua vantagem de ser um modelo flexível que pode ser usado para modelar uma variedade de dados e, é um modelo relativamente fácil de estimar (Nkounkou et al., 2017).

Da equação (2.29) se $w(\cdot) = b, \forall n \in \mathbb{N}$, onde b é uma constante qualquer, então a DPP reduz-se à distribuição de Poisson.

$$p_w(n) = \frac{w(n)p(n)}{\mathbb{E}[w(N)]} = \frac{bp(n)}{\sum_{k=0}^{\infty} bp(k)} = \frac{bp(n)}{b \sum_{k=0}^{\infty} p(k)} = p(n), \quad (2.30)$$

uma vez que $f(k)$ é a função de massa de probabilidade de Poisson e, portanto, a sua soma é 1.

Se a função peso na equação (2.29) for $w(n) = n$ então

$$\mathbb{E}[w(N)] = \sum_{n=0}^{\infty} \frac{n\lambda^n e^{-\lambda}}{n!} = \lambda \sum_{n'=0}^{\infty} \frac{\lambda^{n'} e^{-\lambda}}{n'} = \lambda, \quad (2.31)$$

$$p(n)_w = \frac{n\lambda^n e^{-\lambda}}{n!} \frac{1}{\lambda} = \frac{\lambda^{n-1} e^{-\lambda}}{(n-1)!}, \quad n \geq 1. \quad (2.32)$$

Para calcular a esperança da variável aleatória da DPP, Podemos calcular isso usando a função gera-

dora de momentos $g(z)$. Para a DPP, temos

$$\begin{aligned}
 g(z) &= \sum_{n=1}^{\infty} \frac{\lambda^{n-1} e^{-\lambda}}{(n-1)!} z^n \\
 &= z \sum_{n'=0}^{\infty} \frac{\lambda^{n'} e^{-\lambda}}{(n')!} z^{n'} \\
 &= z \sum_{n'=0}^{\infty} \frac{e^{-\lambda} (\lambda z)^{n'}}{(n')!} = z e^{-\lambda} e^{\lambda z} = z e^{\lambda(z-1)}.
 \end{aligned} \tag{2.33}$$

A esperança da DPP é

$$E[(N_w)] = \dot{g}(1),$$

$$\dot{g}(z) = e^{\lambda(z-1)} + z\lambda e^{\lambda(z-1)}, \tag{2.34}$$

$$E[(N_w)] = \dot{g}(1) = 1 + \lambda. \tag{2.35}$$

A variância da DPP é

$$\text{Var}(N_w) = E[N_w^2] - [E(N_w)]^2 = \ddot{g}(1) + \dot{g}(1) - [\dot{g}(1)]^2, \tag{2.36}$$

$$\ddot{g}(z) = 2\lambda e^{\lambda(z-1)} + z\lambda^2 e^{\lambda(z-1)}, \tag{2.37}$$

então a variância é

$$\begin{aligned}
 \text{Var}(N_w) &= \ddot{g}(1) + \dot{g}(1) - [\dot{g}(1)]^2 \\
 &= 2\lambda + \lambda^2 + 1 + \lambda - (\lambda + 1)^2 \\
 &= 2\lambda + \lambda^2 + 1 + \lambda - \lambda^2 - 2\lambda - 1 \\
 &= \lambda.
 \end{aligned} \tag{2.38}$$

Tabela 2.1: Funções de peso frequentemente usadas

Função Peso	Descrição
Se $w(n) = n$	A DPP resultante é conhecida como distribuição de Poisson “size-based”. Esta distribuição específica foi discutida por Snyder, Neugebauer e Rao (1963) e por Kokonendji, Mizere e Balakrishnan (2008).
Se $w(n; \phi) = (n!)^{1-\alpha}$	A DPP é igual à DCMP discutida anteriormente nesta seção. Esta igualdade tem sido discutida em muitos trabalhos, um dos quais é o de Kokonendji, Mizere e Balakrishnan (2008). Como já foi referido, a DCMP pode ser subdisperso, superdisperso ou equidispersa, dependendo do valor de α .
Se $w(n; \phi) = (n + a)^r$	Em que $a > 0$ é um parâmetro de deslocamento e $r \in \mathbb{R}$ foi discutido em profundidade por Del Castillo e Pérez-Casany (1998). Esta DPP é superedispersa, subdispersa e equidispersa se e somente se $r < 0$, $r > 0$ e $r = 0$, respectivamente. Assim, o parâmetro r pode ser interpretado como o parâmetro de dispersão. Del Castillo e Pérez-Casany (1998) derivaram muitas propriedades da distribuição de Poisson ponderada quando $w(n; \phi) = (n + a)^r$.
se $w(n) = (n + 1)^\delta e^\lambda$	Função de peso da distribuição de Touchard, onde $\delta \in \mathbb{R}$.

Mijburgh (2020) propôs algumas funções de peso, mas não explica nitidamente a essência do seu surgimento, se são relacionados com os dados e em que circunstância devemos usá-las.

Tabela 2.2: Funções de peso

Função Peso	Descrição
$w(n, \phi) = n^{-a}$	Com $a \in \mathbb{N}_1$.
$w(n; \phi) = an^3 + bn^2 + cn$	Com $n \geq 1$, $a, b, c \geq 0$, $\lambda > 0$.
$w(n, \phi) = (n + a)(n - b)^2$	na qual $a, b \geq 0$, $\lambda > 0$.
$w(n, \phi) = (n + a)(n^2 - bn + c)$	O problema com $w(n, \phi) = (n + a)(n^2 - bn + c)$, no entanto, é que introduz restrições adicionais sobre os parâmetros contidos na função peso. Devido ao facto de $n^2 - bn + c$ deve não ter raízes reais, uma restrição adicional é que $b^2 - 4c < 0$. Restrições adicionais de parâmetros são também introduzidas pela expressão da função de massa de probabilidade.
$w(n, \phi) = \frac{b-ac}{n+c} + a$	Com $n \in \mathbb{N}a$, $b > 0$, $\lambda > 0$, $c \in \mathbb{N}_0$

Tabela 2.3: Funções de peso não invertíveis

Função Peso	Razão para a ausência de uma função de massa de
n^a	Embora existam expressões de forma fechada para valores inteiros específicos de a , não existe uma fórmula geral.
$(n + \varepsilon)^{-1}$	É possível obter expressões de forma fechada para esta função de peso. No entanto, as expressões incluem um termo $(-\lambda)^\varepsilon$. Uma vez que $\lambda > 0$ e $\varepsilon \in (0, 1)$, o termo é complexo e, portanto, a função de massa de probabilidade resultante é inválida.
$w(n, \phi) = (n + a)(n - b)^2$	No qual $a, b \geq 0, \lambda > 0$.
$w(n, \phi) = (n + a)(n^2 - bn + c)$	O problema com $w(n, \phi) = (n + a)(n^2 - bn + c)$, no entanto, é que introduz restrições adicionais sobre os parâmetros contidos na função peso. Devido ao facto de $n^2 - bn + c$ deve não ter raízes reais, uma restrição adicional é que $b^2 - 4c < 0$. Restrições adicionais de parâmetros são também introduzidas pela expressão da função de massa de probabilidade.
$(na^3 + bn^2 + cn)^{-1}$	Existem expressões de forma fechada para esta função de peso. No entanto, as expressões, mesmo depois de totalmente simplificadas simplificadas, são extremamente longas. Por exemplo, as expressões da constante de normalização e da função de massa de probabilidade e da função de massa de probabilidade têm, cada uma, mais de meia página de comprimento.
$((n + a)(n - b)^2)^{-1}$	Podem ser derivadas expressões de forma fechada para esta função de peso. No entanto, as expressões incluem muitos termos que levam a contradições nas restrições dos parâmetros. O exemplo mais óbvio é o facto de se exigir que λ seja maior e menor que 0. Consequentemente, esta função peso não conduz a uma função de massa de probabilidade válida.
$((n + a)(n^2 - bn + c))^{-1}$	É possível obter expressões de forma fechada para esta função de peso. No entanto, as expressões incluem termos que levam a contradições nas restrições dos parâmetros. Consequentemente, esta função de ponderação não conduz a uma função de massa de probabilidade válida.
$\binom{m}{n} p^n (1 - p)^{m-n}^{-1}$	Não existe uma expressão geral em forma fechada para esta função de peso não existe. No entanto, se se assumir que o parâmetro m fixo, existe uma expressão. Uma vez que m é o limite superior do domínio desta DPP específica, não é irrealista assumir que praticamente o valor de m poderia ser determinado por reais.
$\binom{m}{n} p^n (1 - p)^{m-n}^{-1} \left(\frac{ab^\alpha}{n^{\alpha+1}}\right)^{-1}$	Embora existam expressões de forma fechada para valores específicos de a e b , não existe uma fórmula geral.

Tabela 2.4: Funções de peso não invertíveis

Função Peso	Razão para a ausência de uma função de massa de
$(I(n \geq a))^{-1}$	A função de ponderação não invertida tem probabilidade 0 quando $n < a$. Em contrapartida, a distribuição de Poisson ponderada invertida (DPPI) é indefinida quando $n < a$. Se o domínio da função for assumido como sendo o mais prático $n \in \{a, a + 1, \dots\}$, a função existe. No entanto, uma vez que $(I(n \geq a))^{-1}$ devolve 1 se for verdadeira, os valores não nulos da função de massa de probabilidade são os mesmos da versão não invertida. O mesmo fato aplica-se às três funções de ponderação de indicadores na versão não invertida.
$(I(n \leq a))^{-1}$	A função de ponderação não invertida tem probabilidade 0 quando $n > b$. Em contrapartida, DPPI é indefinida quando $n > b$.
$(I(n \geq a)I(n \leq a))^{-1}$	A função de ponderação não invertida tem probabilidade 0 quando $n < a$ ou $n > b$. Em contrapartida, a DPPI é indefinida quando $n < a$ ou $n > b$.
$\left(\varepsilon \frac{n!}{e^{-\lambda} \lambda^n I(n=0) + (1-\varepsilon)}\right)^{-1}$	Existe uma expressão em forma fechada para esta função de peso específica. No entanto, a forma da função de massa de probabilidade resultante não é substancialmente afetada pelo valor de ε .

Teorema 4. *Considere a função peso de Poisson da forma exponencial $w(n) = e^{rt(n)}$, $n \in \mathbb{N}$, em que $r \in \mathbb{R}$ e $n \rightarrow t(n)$ é uma função convexa (que pode ou não depender do parâmetro de Poisson original). Se $r > 0$, $r = 0$ e $r < 0$, respectivamente, então o DPP correspondente é superdisperso, subdisperso ou equidisperso, respetivamente (Kokonendji, Mizere e Balakrishnan, 2008).*

Teorema 5.

Sejam $\phi > 0$ e $\theta \in \Theta \subseteq \mathbb{R}$. Então, a função de massa de probabilidade (fmp) é da forma $p(n; \theta, \phi) = \Phi(n; \phi) \exp[n\theta - K(\theta; \phi)]$, $n \in \mathbb{N}$, é uma DPP em relação à distribuição de Poisson padrão com média $\exp(\theta)$. Neste caso, a função de peso de Poisson correspondente é $w(n; \phi) = n! \Phi(n; \phi)$, $n \in \mathbb{N}$ (Kokonendji, Mizere e Balakrishnan, 2008).

Então

$$E_{\theta} [w(N, \phi)] = \sum_{n=0}^{\infty} w(n, \phi) (n!)^{-1} \exp(n\theta - \exp(\theta)) = \sum_{n=0}^{\infty} n! \Phi(n; \phi) (n!)^{-1} \exp(n\theta - \exp(\theta))$$

$$= \sum_{n=0}^{\infty} \Phi(n; \phi) \exp(n\theta - \exp(\theta)) = \sum_{n=0}^{\infty} \Phi(n; \phi) \exp[-\exp(\theta) + K(\theta; \phi)], \quad (2.39)$$

e, por conseguinte, podemos escrever a fmp como

$$\begin{aligned} P_w(n; \theta, \phi) &= \Phi(n, \phi) \exp[n\theta - k(\theta, \phi)] \\ &= \frac{n! \Phi(n, \phi)}{\exp[-\exp(\theta) + k(\theta, \phi)]} \frac{\exp[n\theta - \exp(\theta)]}{n!} \\ &= \frac{w(n, \phi) p(n, \theta)}{\mathbb{E}_\theta[w(N, \phi)]} \\ &= P_w(n, \theta). \end{aligned}$$

Lema 1. *Seja N uma variável aleatória de Poisson com média $\mu > 0$, e seja $w(n) = w(n; \phi)$, $n \in \mathbb{N}$, uma função peso (Kokonendji, Mizere e Balakrishnan, 2008).*

Portanto

$$\text{Var}_\mu(N_w) = \mathbb{E}_\mu(N_w) + \mu^2 \frac{d^2}{d\mu^2} \log \mathbb{E}_\theta[w(N, \phi)]. \quad (2.40)$$

Seja $\theta = \mu$, então, a fmp de N_w é dada por

$$\begin{aligned} P_w(n, \theta) &= \Phi(n, \phi) \exp[n\theta - k(\theta, \phi)] \\ &= \frac{n! \Phi(n, \phi)}{\exp[-\exp(\theta) + k(\theta, \phi)]} \frac{\exp[n\theta - \exp(\theta)]}{n!} \\ &= \frac{w(n, \phi) p(n, \theta)}{\mathbb{E}_\theta[w(N, \phi)]} \\ &= P_w(n, \theta). \end{aligned}$$

Como

$$P_w(n, \theta) = \frac{w(n, \phi)}{\mathbb{E}_\theta[w(N, \phi)]} \frac{\exp[n\theta - \exp(\theta)]}{n!},$$

também podemos expressar $P_w(n, \theta)$ por

$$P_w(n, \theta) = \frac{w(n, \phi)}{n!} (\mathbb{E}_\theta [w(N, \phi)])^{-1} \exp [n\theta - \exp(\theta)]. \quad (2.41)$$

$$P_w(n, \theta) = \frac{w(n, \phi)}{n!} \exp (n\theta - \exp (\theta) - \log \mathbb{E}_\theta [w(N, \phi)]). \quad (2.42)$$

$$P_w(n, \theta) = \frac{w(n, \phi)}{n!} \exp (n\theta - [\exp(\theta) + \log \mathbb{E}_\theta [w(N, \phi)]]). \quad (2.43)$$

$$\mathbb{E}_\mu (N_w) = \exp (\theta) + \frac{d}{d\theta} \log \mathbb{E}_\theta [w(N, \phi)] = \mu + \mu \frac{d}{d\mu} \log \mathbb{E}_\mu [w(N, \phi)]. \quad (2.44)$$

$$\text{Var}_{\mu(\theta)} (N_w) = \frac{d^2}{d\theta^2} k(\theta, \phi). \quad (2.45)$$

$$\mathbb{E}_\mu (N_w) = \exp (\theta) + \frac{d}{d\theta} \log \mathbb{E}_\theta [w(N, \phi)] = \mu + \mu \frac{d}{d\mu} \log \mathbb{E}_\mu [w(N, \phi)]. \quad (2.46)$$

$$\begin{aligned} \text{Var}_{\mu(\theta)} (N_w) &= \exp (\theta) + \frac{d}{d\theta} \left(\frac{d}{d\theta} \log \mathbb{E} [w(N, \phi)] \right) \\ &= \mu + \mu \frac{d}{d\mu} \left(\mu \frac{d}{d\theta} \log \mathbb{E} [w(N, \phi)] \right) \\ &= \mu + \mu \frac{d}{d\mu} \log \mathbb{E} [w(N, \phi)] + \mu^2 \frac{d^2}{d\mu^2} \log \mathbb{E} [w(N, \phi)] \\ &= \mathbb{E}_\mu (N_w) + \mu^2 \frac{d^2}{d\mu^2} \log \mathbb{E}_\theta [w(N, \phi)]. \end{aligned} \quad (2.47)$$

Teorema 6. *Seja N uma variável aleatória de Poisson com média $\mu > 0$, e seja $w(n) = w(n, \phi)$, $n \in \mathbb{N}$ seja uma função de peso que não dependente de μ . Então, a função de peso médio $\mu \rightarrow \mathbb{E} [w(N, \phi)]$ é logconvexo (logcôncavo) se e só se a versão ponderada N_w de N for superdispersa ou subdispersa.*

Log-convexo e log-côncavo são termos usados para descrever propriedades específicas de funções em relação à sua convexidade ou concavidade depois de aplicarmos a função logaritmo natural (\ln). Uma função $f(x)$ é chamada de log-côncava se $\log f(x)$ é uma função côncava. Isso significa que se tomarmos o logaritmo natural de $f(x)$, a função resultante é côncava. Formalmente, $f(x)$ é log-côncava

se para quaisquer x_1 e x_2 no domínio de f e para qualquer $\lambda \in [0, 1]$, temos

$$\log f(\lambda x_1 + (1 - \lambda)x_2) \geq \lambda \log f(x_1) + (1 - \lambda) \log f(x_2)$$

Isso significa que $f(x)$ decresce exponencialmente rápido ou, mais precisamente, que a taxa de decrescimento de $f(x)$ não aumenta.

De fato, utilizando o Lema 1, o índice de Fisher de N_w é dado por:

$$I_\mu(N_w) = \frac{\text{Var}_{\mu(\theta)}(N_w)}{\text{E}_\mu(N_w)} = \frac{\text{E}_\mu(N_w) + \mu^2 \frac{d^2}{d\mu^2} \log \text{E}_\theta[w(N, \phi)]}{\text{E}_\mu(N_w)} = 1 + \frac{\mu^2}{\text{E}_\mu(N_w)} \frac{d^2}{d\mu^2} \log \text{E}_\mu[w(N, \phi)], \quad (2.48)$$

desde que

$$\mu \rightarrow \text{E}[w(N, \phi)] = \sum_{n=0}^{\infty} w(n, \phi) \frac{\mu^n \exp(-\mu)}{n!}.$$

Se $I_\mu(N_w) > 1$ (Superdispersão);

Se $I_\mu(N_w) < 1$ (Subdispersão).

2.2.6 Distribuição de Touchard

Castellares, Lemonte e Moreno-Arenas (2019), introduziram a distribuição discreta de dois parâmetros com base na seguinte expansão em série apresentada em Touchard (1933).

$$T_i(N) = e^n \sum_{k=0}^{\infty} \frac{k^i n^k}{k!}, \quad (2.49)$$

em que $i = 1, 2, 3, \dots$ e $n = 0, 1, 2, \dots$

Por conhecimento da limitação da distribuição de Poisson, foram sugeridas várias generalizações dessa distribuição para tratar dados de contagem em que haja a superdispersão, subdispersão ou excesso de zeros. A distribuição de Touchard (DT) é uma distribuição mais simples e flexível para permitir uma vasta gama de subdispersão, superdispersão e excesso de zeros (Matsushita et al., 2019).

Definição 1. *Seja N uma variável aleatória com valores inteiros não negativos, $n \in \mathbb{N}$, cuja distribuição*

de probabilidade é definida como

$$p_n = P[N = n] = \frac{\lambda^n (n+1)^\delta}{n! \tau(\lambda, \delta)}, \quad (2.50)$$

no qual $\lambda > 0$ e $\delta \in \mathbb{R}$ são os parâmetros da distribuição, e a função $\tau(\lambda, \delta)$

$$\tau(\lambda, \delta) = \sum_{j \in \mathbb{N}} \frac{\lambda^j (j+1)^\delta}{j!}. \quad (2.51)$$

A equação (2.51) é a normalizadora, que está relacionada com polinômios de Touchard (Chrysaphinou, 1985) e com o momento de ordem δ de uma distribuição de Poisson deslocada. Assim, como se constata em (2.50), seja $N \sim \text{Touchard}(\lambda, \delta)$, definida em como uma generalização da distribuição de Poisson já que para $\delta = 0$, $N \sim \text{Poisson}(\lambda)$ (Matsushita et al., 2019).

Resolvendo recursivamente, a equação (2.50) pode ser escrita de seguinte maneira

$$\frac{p_{n+1}}{p_n} = \frac{\lambda^n \lambda (n+2)^\delta}{(n+1)n! \tau(\lambda, \delta)} \frac{n! \tau(\lambda, \delta)}{\lambda^n (n+1)^\delta}.$$

Momentos

O r -ésimo momento de uma variável aleatória de Touchard é uma série polinomial de tipo binomial dada por (Matsushita et al., 2019).

$$\begin{aligned} E(N^r) &= \sum_{n=0}^{\infty} n^r P(N = n) = \sum_{n=0}^{\infty} \frac{n^r (n+1)^\delta \lambda^n}{n! \tau(\lambda, \delta)} \\ &= \frac{1}{\tau(\lambda, \delta)} \sum_{n=0}^{\infty} \frac{n^r (n+1)^\delta \lambda^n}{n!} \\ &= \frac{1}{\tau(\lambda, \delta)} \sum_{j=0}^r \binom{r}{j} (-1)^{r-j} \sum_{n=0}^{\infty} \frac{\lambda^n (n+1)^{\delta+j}}{n!} \\ &= \frac{1}{\tau(\lambda, \delta)} \sum_{j=0}^r \binom{r}{j} (-1)^{r-j} \tau(\lambda, \delta + j) \\ &= \sum_{j=0}^r \binom{r}{j} (-1)^{r-j} \frac{\tau(\lambda, \delta + j)}{\tau(\lambda, \delta)}. \end{aligned} \quad (2.52)$$

E a função geradora de momentos de $N \sim Touchard(\lambda, \delta)$ é dada por

$$\begin{aligned}
 M_N(q) = E[e^{qN}] &= \sum_{n=0}^{\infty} e^{qn} P(N = n) = \sum_{n=0}^{\infty} e^{qn} \frac{\lambda^n (n+1)^\delta}{n! \tau(\lambda, \delta)} \\
 &= \sum_{n=0}^{\infty} \frac{(\lambda e^q)^n (n+1)^\delta}{n! \tau(\lambda, \delta)} \\
 &= \sum_{n=0}^{\infty} \frac{(\lambda e^q)^n (n+1)^\delta}{n!} \frac{1}{\tau(\lambda, \delta)} \\
 &= \frac{\tau(\lambda e^q, \delta)}{\tau(\lambda, \delta)}, \tag{2.53}
 \end{aligned}$$

em que $q \in \mathbb{R}$. Portanto, a DT pertence à família de distribuições de séries de potência de dois parâmetros, a média de N pode ser expressa por Johnson, Kemp e Kotz (2005).

$$\begin{aligned}
 \mu = E[N] &= \sum_{n=0}^{\infty} n \frac{\lambda^n (n+1)^\delta}{n! \tau(\lambda, \delta)} \\
 &= \sum_{n=1}^{\infty} \frac{\lambda^n (n+1)^\delta}{(n-1)! \tau(\lambda, \delta)} \\
 &= \lambda \sum_{n'=0}^{\infty} \frac{\lambda^{n'} (n'+2)^\delta}{n'! \tau(\lambda, \delta)} \\
 &= \lambda \sum_{n'=0}^{\infty} \frac{(n'+2)^\delta}{(n'+1)^\delta} \frac{\lambda^{n'} (n'+1)^\delta}{n'! \tau(\lambda, \delta)} \\
 &= \lambda \sum_{n'=0}^{\infty} \left(\frac{n'+2}{n'+1} \right)^\delta \cdot p_{n'} \\
 &= \lambda E \left[\left(\frac{N+2}{N+1} \right)^\delta \right], \tag{2.54}
 \end{aligned}$$

$$\mu = E[N] = \frac{\tau(\lambda, \delta+1)}{\tau(\lambda, \delta)} - 1 = \lambda \cdot E \left[\left(\frac{N+2}{N+1} \right)^\delta \right]. \tag{2.55}$$

Agora vamos calcular a variância, mas antes temos que encontrar $E[N^2]$

$$\begin{aligned} E[N^2] &= \binom{2}{0} (-1)^2 \frac{\tau(\lambda, \delta)}{\tau(\lambda, \delta)} + \binom{2}{1} (-1) \frac{\tau(\lambda, \delta + 1)}{\tau(\lambda, \delta)} + \binom{2}{2} (-1)^0 \frac{\tau(\lambda, \delta + 2)}{\tau(\lambda, \delta)} \\ &= 1 - 2 \frac{\tau(\lambda, \delta + 1)}{\tau(\lambda, \delta)} + \frac{\tau(\lambda, \delta + 2)}{\tau(\lambda, \delta)}, \end{aligned} \quad (2.56)$$

então a variância é dada por

$$\begin{aligned} \sigma^2 &= \text{Var}(N) = E(N^2) - [E(N)]^2 \\ &= 1 - 2 \frac{\tau(\lambda, \delta + 1)}{\tau(\lambda, \delta)} + \frac{\tau(\lambda, \delta + 2)}{\tau(\lambda, \delta)} - \left[\frac{\tau(\lambda, \delta + 1)}{\tau(\lambda, \delta)} - 1 \right]^2 \\ &= \frac{\tau(\lambda, \delta + 2)}{\tau(\lambda, \delta)} - \left[\frac{\tau(\lambda, \delta + 1)}{\tau(\lambda, \delta)} \right]^2. \end{aligned} \quad (2.57)$$

Ou ainda, podemos expressar em termos de μ como

$$\begin{aligned} \sigma^2 &= \text{Var}(N) = E(N^2) - [E(N)]^2 \\ &= 1 - 2 \frac{\tau(\lambda, \delta + 1)}{\tau(\lambda, \delta)} + \frac{\tau(\lambda, \delta + 2)}{\tau(\lambda, \delta)} - \mu^2 \\ &= 1 - 2(\mu + 1) + \frac{\tau(\lambda, \delta + 2)}{\tau(\lambda, \delta)} - \mu^2 \\ &= \frac{\tau(\lambda, \delta + 2)}{\tau(\lambda, \delta)} - (\mu + 1)^2 \\ &= \frac{\tau(\lambda, \delta + 2)\tau(\lambda, \delta + 1)}{\tau(\lambda, \delta)\tau(\lambda, \delta + 1)} - (\mu + 1)^2 \\ &= \frac{\tau(\lambda, \delta + 2)}{\tau(\lambda, \delta + 1)} \cdot (\mu + 1) - (\mu + 1)^2 \\ &= (\mu + 1) \cdot \left[\frac{\tau(\lambda, \delta + 2)}{\tau(\lambda, \delta + 1)} \right]. \end{aligned} \quad (2.58)$$

Da equação (2.54), observa-se que $\mu > \lambda$, se $\delta > 0$; e $\mu < \lambda$, se $\delta < 0$. Para avaliar a dispersão, considere-se o rácio $r = \sigma^2/\mu$, que pode ser expresso por

$$r = \frac{\lambda \cdot E \left[(N + 1) \left(\frac{N+2}{N+1} \right)^\delta \right] - \mu^2}{\lambda \cdot E \left[\left(\frac{N+2}{N+1} \right)^\delta \right]}$$

$$\begin{aligned}
&= \frac{E \left[(N + 1) \left(\frac{N+2}{N+1} \right)^\delta \right]}{E \left[\left(\frac{N+2}{N+1} \right)^\delta \right]} - \frac{\mu^2}{\lambda E \left[\left(\frac{N+2}{N+1} \right)^\delta \right]} \\
&= \frac{E \left[(N + 1) \left(\frac{N+2}{N+1} \right)^\delta \right]}{E \left[\left(\frac{N+2}{N+1} \right)^\delta \right]} - \mu. \tag{2.59}
\end{aligned}$$

A equação (2.59) descreve o rácio de dispersão $r = \frac{\sigma^2}{\mu}$, que é uma medida da variabilidade em relação à média de uma distribuição. Essa fórmula está relacionada com a Distribuição de Touchard, onde δ atua como um parâmetro que controla a dispersão. Vamos analisar o comportamento de r em função de δ :

- Quando $\delta = 0$, a equação reduz-se à distribuição de Poisson, onde o valor esperado (μ) é igual à variância (σ^2). Nesse caso, o rácio $r = 1$, o que indica que a dispersão é equidispersa, ou seja, a variância é igual à média, como é característico da distribuição de Poisson.
- Para $\delta > 0$, o termo $\left(\frac{N+2}{N+1} \right)^\delta$ resulta em valores maiores para $E \left[(N + 1) \left(\frac{N+2}{N+1} \right)^\delta \right]$ no numerador, aumentando a diferença entre a média e a variância. No entanto, o efeito é que a variância σ^2 se torna menor do que a média μ , resultando em $r < 1$, o que caracteriza subdispersão. Isso significa que os dados são menos dispersos do que o esperado em uma distribuição de Poisson.
- Para $\delta < 0$, o termo $\left(\frac{N+2}{N+1} \right)^\delta$ reduz-se, diminuindo o valor esperado no numerador. Nesse caso, a variância σ^2 se torna maior do que a média μ , o que resulta em $r > 1$, caracterizando superdispersão. Neste cenário, os dados são mais dispersos, com maior variabilidade em relação à média, em comparação com a distribuição de Poisson.

Portanto, o parâmetro δ desempenha um papel crucial na definição da dispersão da Distribuição de Touchard. A relação entre δ e r permite modelar diferentes comportamentos de dispersão: $\delta > 0$ para subdispersão, $\delta = 0$ para equidispersão, e $\delta < 0$ para superdispersão.

Capítulo 3

Processo de Poisson

Neste capítulo, serão apresentados os fundamentos do Processo de Poisson, incluindo suas principais definições e características matemáticas. Iniciaremos com uma introdução às propriedades básicas do processo, seguido pela definição formal do mesmo, enfatizando sua aplicabilidade em eventos que ocorrem de forma independente e aleatória ao longo do tempo. Em seguida, abordaremos a função de probabilidade para o número de eventos em um intervalo de tempo de comprimento específico, estabelecendo as bases para o cálculo de probabilidades em processos de contagem. Por fim, exploraremos as equações de diferença-diferencial associadas ao Processo de Poisson, que descrevem a evolução temporal da distribuição de probabilidade e são fundamentais para o entendimento e a modelagem de processos contínuos e discretos. Essa estrutura teórica permitirá a compreensão e o desenvolvimento de extensões do Processo de Poisson que melhor capturam a natureza dos dados considerados neste estudo.

3.1 Processo de Poisson Homogêneo

Definição 2. Um processo estocástico $\{N(t), t \in T\}$ é uma coleção de variáveis aleatórias. Ou seja, para cada t no conjunto T , $N(t)$ é uma variável aleatória. Se T for um conjunto contável, o processo estocástico diz-se “em tempo discreto”, e se T for um conjunto contínuo, diz-se “em tempo contínuo”. Se $N(t)$ só pode assumir valores inteiros, diz-se que o processo estocástico é um processo estocástico “de valor inteiro”. Se $N(t)$ puder assumir qualquer valor real, o processo é designado por processo estocástico de “valor real.” (Ross, 2014).

Definição 3. Um processo de contagem $\{N(t), t \geq 0\}$, conforme a definição (2), é chamado de processo de Poisson com taxa $\lambda > 0$ se

1. $N(0) = 0$.
2. O processo tem incrementos independentes.
3. $N(t_1) - N(t_0), N(t_2) - N(t_1), \dots, N(t_n) - N(t_{n-1})$ são variáveis aleatórias independentes onde $t_0 < t_1 < \dots < t_n$;
4. O número de eventos em qualquer intervalo de comprimento t é distribuído segundo uma Poisson com média λt . Em outras palavras, para $0 < s < t$, temos que

$$P[N(t+s) - N(s) = n] = \frac{e^{-\lambda t} (\lambda t)^n}{n!}, \quad n = 0, 1, 2, \dots \quad (3.1)$$

Definição 4. Um processo de contagem $\{N(t), t \geq 0\}$ diz-se ser um processo de Poisson com taxa $\lambda > 0$ se

$$N(t) = \sum_{j=1}^{\infty} I_{[0,t]}(U_j), \quad t \geq 0,$$

na qual, a sequência (U_j) é i.i.d. e os incrementos t_1, t_2, \dots , cada um dos quais segue uma distribuição exponencial. Os U_j são variáveis aleatórias uniformes padrão independentes, que também são independentes de N e são conhecidos como épocas de chegada e o intervalo $[0, t]$ são os tempos de inter-chegada (Balakrishnan e Kozubowski, 2008).

Definição 5. Diz-se que um processo estocástico $\{U(t), t \geq 0\}$ é um processo de Poisson composto se puder ser representado, para $t \geq 0$, por

$$U(t) = \sum_{i=1}^{N(t)} U_i,$$

em que $\{N(t), t \geq 0\}$ é um processo de Poisson e U_1, U_2, \dots é uma família de variáveis aleatórias independentes e identicamente distribuídas que é independente do processo $\{N(t), t \geq 0\}$ (Balakrishnan e Kozubowski, 2008).

3.2 Função de Probabilidade para um Intervalo de Tempo de Comprimento

Considere uma contagem aleatória N por unidade de tempo ou espaço como, por exemplo, o fluxo de veículos (número de veículos por unidade de tempo), a incidência mensal de certa doença (número mensal de casos novos da doença) e a densidade populacional (número de pessoas por metro quadrado). Como essa contagem depende de uma exposição ao intervalo de tempo (ou espaço), escrevemos a variável aleatória N como $N(t)$, na qual $t > 0$ representa o comprimento do intervalo de tempo $[0; t]$ que constitui a janela de exposição temporal do experimento aleatório. Se esse comprimento for nulo, é razoável estabelecer a condição inicial $N(0) \equiv 0$.

Assim, por exemplo, $N(t_1)$ expressa o número de ocorrências registradas no intervalo de tempo $[0; t_1]$. Se $t_2 > t_1$, por se tratar de contagens, $N(t_2) \geq N(t_1)$, e a diferença $N(t_2) - N(t_1)$ representa a contagem registrada no intervalo de tempo $[t_1; t_2]$, cujo comprimento é $t_2 - t_1$.

Agora, suponha que $N(t_2 - t_1) = N(t_2) - N(t_1)$, ou seja, a contagem N depende apenas do intervalo de tempo, mas não depende da escolha do momento inicial. Além disso, considere que as contagens realizadas em intervalos de tempo não sobrepostos sejam independentes. Por exemplo, considerando os instantes $t_1 < t_2 < t_3$, as contagens $N(t_2) - N(t_1)$ e $N(t_3) - N(t_2)$ são independentes.

Tome agora um intervalo de tempo de comprimento $\Delta t > 0$ suficientemente pequeno, de modo que seja praticamente improvável a ocorrência de 2 ou mais registros, ou seja, $P[N(\Delta t) \geq 2] \approx 0$. Nesse mesmo intervalo, assuma que a probabilidade de ocorrência de um caso seja proporcional ao comprimento do intervalo de tempo, ou seja, $P[N(\Delta t) = 1] \approx \lambda \Delta t$, em que $\lambda > 0$ é uma constante de proporcionalidade. Consequentemente, $P[N(\Delta t) = 0] \approx 1 - \lambda \Delta t$. Portanto, $N(\Delta t)$ segue, aproximadamente, uma distribuição de Bernoulli.

3.3 Equações de Diferença-Diferencial

Uma **equação de diferença-diferencial** é uma equação matemática que combina características de **equações diferenciais** e **equações de diferenças**. Essas equações envolvem tanto derivadas (relacionadas a variações contínuas, típicas de equações diferenciais) quanto diferenças (relacionadas a variações discretas, típicas de equações de diferenças). Elas são utilizadas em situações onde o comportamento de

um sistema depende tanto de aspectos contínuos quanto discretos, ou seja, quando as mudanças ocorrem ao longo de intervalos contínuos e discretos de tempo ou espaço (ELAYDI e ZIEGLER, 2004).

Estrutura

Uma equação de diferença-diferencial pode ser vista como uma generalização das equações diferenciais e de diferenças. Ela possui termos de derivadas em relação ao tempo (ou outra variável contínua) e também diferenças em relação a valores passados ou discretos de uma função. Um exemplo simples de uma equação de diferença-diferencial seria:

$$\frac{dy(t)}{dt} = ay(t) + by(t - n)$$

Aqui, $\frac{dy(t)}{dt}$ é uma derivada com relação ao tempo contínuo t , enquanto $y(t - n)$ representa a função com atraso discreto de n unidades de tempo. Esse tipo de equação modela sistemas com **atrasos** e é usada em diversos campos, como controle de sistemas, biologia (modelagem de populações), economia e engenharia (ELAYDI e ZIEGLER, 2004).

O objetivo é deduzir a expressão da função de probabilidade para um intervalo de tempo de comprimento t qualquer, $p_n(t) = P[N(t) = n]$ para $n \in \mathbb{N}$. A partir das condições estabelecidas nesse processo de contagem, iniciando com o caso $n = 0$, pode-se escrever

$$\begin{aligned} p_0(t + \Delta t) &= P[N(t + \Delta t) = 0] = P[N(t + \Delta t) - N(t) + N(t) = 0] \\ &= P[N(t + \Delta t) - N(t) = 0, N(t) = 0] \\ &= P[N(t + \Delta t) - N(t) = 0] P[N(t) = 0] \\ &= P[N(\Delta t) = 0] p_0(t) \\ &= (1 - \lambda \Delta t) p_0(t), \end{aligned} \tag{3.2a}$$

de modo que

$$\frac{p_0(t + \Delta t) - p_0(t)}{\Delta t} = -\lambda p_0(t). \tag{3.2b}$$

Na equação (3.2b), fazendo $\Delta t \rightarrow 0$, podemos escrever

$$\begin{aligned}\dot{p}_0(t) &= \frac{d}{dt}p_0(t) \\ &= \lim_{\Delta t \rightarrow 0} \frac{p_0(t + \Delta t) - p_0(t)}{\Delta t} = -\lambda p_0(t).\end{aligned}$$

Obtendo $\frac{\dot{p}_0(t)}{p_0(t)} = -\lambda$, ou seja,

$$\frac{d}{dt} \ln p_0(t) = -\lambda. \quad (3.2c)$$

Integrando ambos lados da equação (3.2c), obtem-se $\ln p_0(t) = -\lambda t$, o que permite concluir que

$$p_0(t) = e^{-\lambda t}. \quad (3.2)$$

De modo análogo, para $n = 1$,

$$\begin{aligned}p_1(t + \Delta t) &= P[N(t + \Delta t) - N(t) + N(t) = 1] \\ &= P[N(t + \Delta t) - N(t) = 1]p_0(t) + P[N(t + \Delta t) - N(t) = 0]p_1(t) \\ &= P[N(\Delta t) = 1]p_0(t) + P[N(\Delta t) = 0]p_1(t) \\ &= \lambda \Delta t p_0(t) + (1 - \lambda \Delta t)p_1(t).\end{aligned} \quad (3.3)$$

Agora,

$$\frac{p_1(t + \Delta t) - p_1(t)}{\Delta t} = \lambda[p_0(t) - p_1(t)], \quad (3.3a)$$

cujos limites para $\Delta t \rightarrow 0$ é

$$\begin{aligned}\dot{p}_1(t) &= \frac{d}{dt}p_1(t) \\ &= \lim_{\Delta t \rightarrow 0} \frac{p_1(t + \Delta t) - p_1(t)}{\Delta t} \\ &= \lambda[p_0(t) - p_1(t)],\end{aligned}$$

ou seja, obtemos uma equação diferencial na forma

$$\dot{p}_1(t) = -\lambda p_1(t) + \lambda e^{-\lambda t}, \quad (3.3b)$$

observando a solução pode ser tal que $p_1(t) \approx e^{-\lambda t}$, pode-se tentar uma solução na forma

$$p_1(t) = a(t) e^{-\lambda t}. \quad (3.3c)$$

Da equação (3.3c) é

$$\dot{p}_1(t) = \dot{a}(t) e^{-\lambda t} - \lambda a(t) e^{-\lambda t} = \dot{a}(t) e^{-\lambda t} - \lambda p_1(t). \quad (3.3d)$$

Igualando a equação (3.3d) com a equação (3.3b), tem-se $\dot{a}(t) e^{-\lambda t} - \lambda p_1(t) = -\lambda p_1(t) + \lambda e^{-\lambda t}$, ou seja,

$$\dot{a}(t) = \lambda,$$

cujas solução é $a(t) = \lambda t$. Da equação (3.3d), obtemos

$$p_1(t) = \lambda t e^{-\lambda t}. \quad (3.4)$$

Para o caso geral $n \geq 1$, tem-se

$$\begin{aligned} p_n(t + \Delta t) &= P[N(t + \Delta t) - N(t) + N(t) = n] \\ &= P[N(t + \Delta t) - N(t) = 1, N(t) = n - 1] + P[N(t + \Delta t) - N(t) = 0, N(t) = n] \\ &= P[N(t + \Delta t) - N(t) = 1]p_{n-1}(t) + P[N(t + \Delta t) - N(t) = 0]p_n(t) \\ &= P[N(\Delta t) = 1]p_{n-1}(t) + P[N(\Delta t) = 0]p_n(t) \\ &= \lambda \Delta t p_{n-1}(t) + (1 - \lambda \Delta t)p_n(t). \end{aligned} \quad (3.5)$$

Nesse caso,

$$\frac{p_n(t + \Delta t) - p_n(t)}{\Delta t} = \lambda[p_{n-1}(t) - p_n(t)], \quad (3.5a)$$

de modo que

$$\begin{aligned} \dot{p}_n(t) &= \frac{d}{dt}p_n(t) \\ &= \lim_{\Delta t \rightarrow 0} \frac{p_n(t + \Delta t) - p_n(t)}{\Delta t} = \lambda[p_{n-1}(t) - p_n(t)]. \end{aligned}$$

Novamente, considerando uma solução geral na forma

$$p_n(t) = a_n(t) e^{-\lambda t}, \quad (3.5b)$$

observe que encontra-se uma forma recursiva $\dot{a}_n(t) = \lambda a_{n-1}(t)$. Aplicando-a sucessivamente para $n = 1$ e 2 vemos que $\dot{a}_1(t) = \lambda$ e $\dot{a}_2(t) = \lambda^2 t$, cujas soluções são $a_1(t) = \lambda t$ e $a_2(t) = \frac{(\lambda t)^2}{2}$. Para o caso geral, podemos concluir que $a_n(t) = \frac{(\lambda t)^n}{n!}$. Assim, finalmente, encontramos a distribuição de Poisson dada por

$$P[N(t) = n] = \frac{(\lambda t)^n e^{-\lambda t}}{n!}. \quad (3.6)$$

Capítulo 4

Processo de Poisson Ponderado

O processo de Poisson ponderado (PrPP) é uma extensão do processo de Poisson clássico, que descreve a chegada de eventos ao longo do tempo. Neste capítulo, o objetivo é transformar um processo de Poisson clássico no intervalo $[0, 1]$ com intensidade λ , $\{N(t), t \in [0, 1]\}$, em um processo ponderado, $\{N_w(t), t \in [0, 1]\}$, onde $w = w_t$ é uma família de funções peso indexadas por $t \in [0, 1]$ (Balakrishnan e Kozubowski, 2008). Os responsáveis na expansão da ideia de DPP à PrPP são Balakrishnan e Kozubowski (2008) e, derivaram muitas das propriedades estatísticas associadas a PrPP. Para o caso do processo de Poisson, temos

$$N(t) = \sum_{j=1}^N I_{[0,t]}(U_j), \quad t \in [0, 1], \quad (4.1)$$

em que $N(t)$ é um processo de Poisson com a taxa $\lambda > 0$, $\{N(t), t \in [0, 1]\}$. N é uma variável aleatória de Poisson com a taxa $\lambda > 0$ e U_j são variáveis aleatórias uniformes padrão independentes, que também são independentes de N .

Definição 6. $N_w(t)$ é dito ser um PrPP com intensidade λ e função peso $w(\cdot)$, no intervalo $[0, 1]$, denotado por $\{N_w(t), t \in [0, 1]\}$, se

$$N_w(t) = \sum_{j=1}^{N_w} I_{[0,t]}(U_j), \quad t \in [0, 1], \quad (4.2)$$

em que N_w é uma variável aleatória de Poisson ponderada com taxa λ e função peso $w(\cdot)$, e U_j são variáveis aleatórias uniformes padrão independentes entre si e de N_w .

Na definição (6) as sequências $I_{[0,t]}(U_j)$ mencionadas são independentes e identicamente distribuídas, e se $t \in [0, 1]$, cada $I_{[0,t]}(U_j)$ é uma variável aleatória de Bernoulli com parâmetro t . Dessa forma, o processo estocástico observado em um dado momento pode, de fato, ser uma variável aleatória composta, cujas propriedades são bem conhecidas na literatura (Balakrishnan e Kozubowski, 2008).

De acordo com o Balakrishnan e Kozubowski (2008) sobre PrPP, foram alcançadas os seguintes resultados

Lema 2. Para cada $t > 0$, a variável aleatória $N_w(t)$ tem uma DPP com parâmetro de taxa λt , e função de peso

$$w_t(k) = \sum_{n=0}^{\infty} \frac{[(1-t)\lambda]^n}{n!} w(n+k), \quad k = 0, 1, 2, \dots \quad (4.3)$$

em que t é incluído no subscrito da função peso para indicar a dependência da função em t .

Sua demonstração pode-se encontrar em Balakrishnan e Kozubowski (2008).

A dispersão de todo o processo estocástico é determinada pela dispersão da variável aleatória de Poisson ponderada relacionada. Assim, para avaliar a dispersão de todo o processo equação (4.2), que é determinada pela natureza da dispersão da variável N_w , que pode ser expresso conforme Proposição 2, a seguir

Proposição 2. Para qualquer $w \in W$, diz-se que o processo dado na equação (4.2) é um PrPP com parâmetro $\lambda > 0$ e função peso w , denotada por $PrPP_\lambda(w)$.

$$\frac{\text{Var}[N_w(t)]}{\text{E}[N_w(t)]} = 1 + t \left(\frac{\text{Var}(N_w)}{\text{E}(N_w)} - 1 \right), \quad 0 \leq s < t \leq 1. \quad (4.4)$$

Se a dispersão da variável aleatória de Poisson ponderada for conhecida, ela também será conhecida para o processo estocástico associado para cada valor específico de t .

Teorema 7. (Propriedades de processos de Poisson ponderados) Suponha-se que $N_w(t)$ é um PrPP, e que N_w é uma variável aleatória de Poisson ponderada, utilizando as propriedades das distribuições

compostas, segue-se que, o valor esperado do processo é dado por

$$\mathbb{E}[N_w(t)] = \mathbb{E}(N_w) \mathbb{E}(I_{[0,t]}[U_j]) = \mathbb{E}(N_w)t. \quad (4.5)$$

A variância do processo é dada por

$$\begin{aligned} \text{Var}[N_w(t)] &= \mathbb{E}(N_w) \text{Var}(I_{[0,t]}[U_j]) + \text{Var}(N_w) \mathbb{E}(I_{[0,t]}[U_j])^2 \\ &= \mathbb{E}(N_w)t(1-t) + \text{Var}(N_w)t^2 \end{aligned} \quad (4.6)$$

A função geradora de probabilidade do processo é dada por

$$\begin{aligned} G_{N_w(t)}(z) &= G_{N_w}(G_{(I_{[0,t]}[U_j])}(z)) \\ &= G_{N_w}(1-t+tz). \end{aligned} \quad (4.7)$$

A função geradora de probabilidade de um incremento do processo é dada por

$$\begin{aligned} G_{N_w(t)-N_w(s)}(z) &= G_{N_w}(G_{(I_{[s,t]}[U_j])}(z)) \\ &= G_{N_w}(1-(t-s)+(t-s)z). \end{aligned} \quad (4.8)$$

A função de massa de probabilidade conjunta do processo é dada por

$$P(N_w(s) = k, N_w(t) = n) = \binom{n}{k} \left(\frac{s}{t}\right)^k \left(1 - \frac{s}{t}\right)^{n-k} P(N_w = n). \quad (4.9)$$

A covariância do processo é dada por

$$\begin{aligned} \text{Cov}(N_w(s), N_w(t)) &= st \text{Var}(N_w) + s(1-t) \mathbb{E}(N_w) \\ &= s(t \text{Var}(N_w) + (1-t) \mathbb{E}(N_w)). \end{aligned} \quad (4.10)$$

A covariância entre dois incrementos independentes do processo é dada por

$$\begin{aligned} \text{Cov}(N_w(s), N_w(t) - N_w(s)) &= \text{Cov}(N_w(s), N_w(t)) - \text{Var}(N_w(s)) \\ &= s(t - s)(\text{Var}(N_w) - E(N_w)). \end{aligned} \quad (4.11)$$

4.1 Função de Peso Polinomial

A função peso $w(n) = n$, é uma função de peso comumente utilizada, esta função é expandida numa família de várias funções diferentes baseadas em polinômios (Mijburgh, 2020).

4.1.1 $w(n) = n$

Teorema 8. Se a equação (4.3) do Lema 2, usada na função de massa de probabilidade de Poisson ponderada for escolhida como $w(n; \phi) = n$ então

$$w_t(n) = e^{\lambda(1-t)}(n + \lambda(1 - t)). \quad (4.12)$$

A distribuição de probabilidade $p_{w,t}(n)$ é dada por

$$P_{w,t}(n) = \frac{e^{-\lambda t}(\lambda t)^n}{n!} \cdot \frac{(n + \lambda(1 - t))}{\lambda} = \frac{e^{-\lambda t}(\lambda t)^n}{n!} \left(\frac{n}{\lambda} + 1 - t \right). \quad (4.13)$$

A função geradora de probabilidade do processo é dada por

$$\begin{aligned} g_{N_w(t)}(z) &= G_{N_w}(1 - t + tz) \\ &= e^{-\lambda t(z-1)}(1 + s(1 + t(z - 1))). \end{aligned} \quad (4.14)$$

A expectativa $E[N_w(t)]$ é

$$E[N_w(t)] = t(\lambda + 1). \quad (4.15)$$

A variância do processo é dada por

$$\begin{aligned}\text{Var}[N_w(t)] &= \text{E}(N_w)t(1-t) + \text{Var}(N_w)t^2 \\ &= (\lambda + 1)t(1-t) + \lambda t^2 \\ &= t(\lambda + 1 - t).\end{aligned}\tag{4.16}$$

A função geradora de probabilidade de um incremento do processo é dada por

$$g_{t-s}(z) = e^{\lambda(t-s)(z-1)}(1 + s(1-z) + t(z-1)).\tag{4.17}$$

A função de massa de probabilidade conjunta do processo é dada por

$$\begin{aligned}\text{P}(N_w(s) = k, N_w(t) = n) &= p_{w,(s,t)}(a, b) \\ &= \binom{n}{k} \left(\frac{s}{t}\right)^k \left(1 - \frac{s}{t}\right)^{n-k} \text{P}(N_w = n) \\ &= \binom{b}{a} \frac{e^{-\lambda t} \left(1 - \frac{s}{t}\right)^{b-a} \left(\frac{s}{t}\right)^a t^b \lambda^{b-1} (b + \lambda(1-t))}{b!}.\end{aligned}\tag{4.18}$$

A covariância do processo é dada por

$$\text{Cov}_{s,t} = s(t \text{Var}(N_w) + (1-t) \text{E}(N_w)) = s(1-t + \lambda).\tag{4.19}$$

A covariância entre dois incrementos independentes do processo é dada por

$$\text{Cov}_{s,t-s} = s(t-s)(\text{Var}(N_w) - \text{E}(N_w)) = s(s-t).\tag{4.20}$$

Na (4.18), os termos a e b são substituições para as variáveis k e n , respectivamente. A equação (4.18) representa a probabilidade conjunta de observar a eventos até o tempo s e b eventos até o tempo t em um processo de Poisson. Estes termos são essenciais para dividir o intervalo $[0, t]$ em duas partes: $[0, s]$ e $(s, t]$, permitindo o cálculo da probabilidade conjunta de eventos ocorrendo nesses subintervalos em um processo de Poisson.

Capítulo 5

O Processo de Touchard

Para mais informação sobre distribuição de Touchard ver em Matsushita e Oliveira (2016).

Como viu-se anteriormente, um processo de Poisson descreve as contagens de um acontecimento de interesse por unidade de tempo se a probabilidade de um acontecimento ocorrer numa qualquer unidade infinitesimal de tempo for independente da probabilidade de acontecer em qualquer outra unidade de tempo. No entanto, quando se analisa dados de contagem, é frequente encontrarmos discrepâncias entre as frequências observadas e as esperadas segundo a lei de Poisson. Vamos agora mostrar como a distribuição de Poisson pode ser modificada de modo a ser adequada, análogo a seção 3.3 do Capítulo 3.

Seja $N(t)$ uma variável aleatória Touchard que toma valores inteiros não negativos num intervalo de tempo $(0, t]$ com distribuição de probabilidade dada por

$$p_n(t) = P[N(t) = n] = \frac{\lambda_t^n (n+1)^{\delta_t}}{n! \tau(\lambda_t, \delta_t)}, \quad (5.1)$$

com condições iniciais $P[N(0) = 0] = 1$, $\lambda_0 = 0$ e $\tau(\lambda_0, \delta_0) = 1$. Para $t > 0$, e a função

$$\tau(\lambda_t, \delta_t) = \sum_{j=0}^{\infty} \frac{\lambda_t^j (j+1)^{\delta_t}}{j!}, \quad (5.2)$$

que normaliza a equação (5.1), é conhecido como o polinómio de Touchard conforme apresentado por Matsushita et al. (2019), Andrade et al. (2021) e Ho et al. (2021). De uma forma simples, pode-se observar que, se $\lambda_t = \lambda t$ e $\delta_t = 0$, então $N(t)$ segue um processo de Poisson com taxa λ , de modo que

a equação (5.1) pode ser considerada uma generalização deste processo.

5.1 Derivadas Temporais

Considere $\lambda_t > 0$ e $\delta_t \in \mathbb{R}$ são funções diferenciáveis do tempo t . A derivada do polinômio conforme na equação (5.1) em relação ao tempo t é

$$\begin{aligned}
 \dot{\tau}(\lambda_t, \delta_t) &= \frac{d}{dt} \tau(\lambda_t, \delta_t) = \frac{d}{dt} \left(\sum_{n=0}^{\infty} \frac{\lambda_t^n (n+1)^{\delta_t}}{n!} \right) = \sum_{n=1}^{\infty} \frac{n \lambda_t^{n-1} \dot{\lambda}_t (n+1)^{\delta_t} + \lambda_t^n (n+1)^{\delta_t} \dot{\delta}_t \ln(n+1)}{n!} \\
 &= \dot{\lambda}_t \sum_{n=1}^{\infty} \frac{n \lambda_t^{n-1} (n+1)^{\delta_t}}{n!} + \dot{\delta}_t \sum_{n=1}^{\infty} \frac{\lambda_t^n (n+1)^{\delta_t} \ln(n+1)}{n!} \\
 &= \frac{\dot{\lambda}_t}{\lambda_t} \sum_{n=1}^{\infty} \frac{\lambda_t^n (n+1)^{\delta_t} [(n+1) - 1]}{n!} \\
 &\quad + \dot{\delta}_t \sum_{n=1}^{\infty} \frac{\lambda_t^n (n+1)^{\delta_t} \ln(n+1)}{n!} \\
 &= \frac{\dot{\lambda}_t}{\lambda_t} \left(\sum_{n=1}^{\infty} \frac{\lambda_t^n (n+1)^{\delta_t+1}}{n!} - \sum_{n=1}^{\infty} \frac{\lambda_t^n (n+1)^{\delta_t}}{n!} \right) \\
 &\quad + \dot{\delta}_t \sum_{n=1}^{\infty} \frac{\lambda_t^n (n+1)^{\delta_t} \ln(n+1)}{n!} \\
 &= \frac{\dot{\lambda}_t}{\lambda_t} [\tau(\lambda_t, \delta_t + 1) - \tau(\lambda_t, \delta_t)] + \dot{\delta}_t \tau^*(\lambda_t, \delta_t),
 \end{aligned} \tag{5.3}$$

em que

$$\tau^*(\lambda_t, \delta_t) = \sum_{n=1}^{\infty} \frac{\lambda_t^n (n+1)^{\delta_t} \ln(n+1)}{n!}.$$

Analogamente, de acordo com o Capítulo 2 da equação (2.55) em que μ_t é expressa por

$$\mu_t = \frac{\tau(\lambda_t, \delta_t + 1)}{\tau(\lambda_t, \delta_t)} - 1. \tag{5.4}$$

Por conseguinte,

$$\begin{aligned}
 \dot{\mu}_t &= \frac{d}{dt} \mu_t = \frac{d}{dt} \sum_{n \geq 0} n p_n(t) = \sum_{n \geq 0} n \dot{p}_n(t) \\
 &= \sum_{n \geq 0} n \left\{ (n - \mu_t) \frac{\dot{\lambda}_t}{\lambda_t} + \dot{\delta}_t (\ln(n+1) - \mathbb{E}[\ln(N(t)+1)]) \right\} p_n(t) \\
 &= \sum_{n \geq 0} \left\{ (n^2 - n\mu_t) \frac{\dot{\lambda}_t}{\lambda_t} + \dot{\delta}_t (n \ln(n+1) - n \mathbb{E}[\ln(N(t)+1)]) \right\} p_n(t) \\
 &= \frac{\dot{\lambda}_t}{\lambda_t} \sigma_t^2 + \dot{\delta}_t \gamma_t,
 \end{aligned} \tag{5.5}$$

em que σ_t^2 conforme na equação (5.18) é

$$\sigma_t^2 = \frac{\tau(\lambda_t, \delta_t + 2)}{\tau(\lambda_t, \delta_t)} - (\mu_t + 1)^2,$$

e

$$\gamma_t = \text{Cov} [N(t), \ln(N(t) + 1)], \tag{5.6}$$

na qual $\dot{\lambda}_t = d\lambda_t/dt$ e $\dot{\delta}_t = d\delta_t/dt$. Neste ponto, por se tratar de um processo de contagens, cabe observar que (5.5) só faz sentido se $\dot{\mu}_t > 0$. Deste modo, a derivada temporal da equação (5.1), para $n \geq 1$, é

$$\begin{aligned}
 \dot{p}_n(t) &= \frac{d}{dt} (p_n(t)) = \frac{d}{dt} \left(\frac{\lambda_t^n (n+1)^{\delta_t}}{n! \tau(\lambda_t, \delta_t)} \right) \\
 &= \frac{n \lambda_t^{n-1} \dot{\lambda}_t (n+1)^{\delta_t} + \lambda_t^n (n+1)^{\delta_t} \dot{\delta}_t \ln(n+1)}{n! \tau(\lambda_t, \delta_t)} - \frac{\lambda_t^n (n+1)^{\delta_t} \dot{\tau}(\lambda_t, \delta_t)}{n! \tau^2(\lambda_t, \delta_t)} \\
 &= \frac{n \lambda_t^{n-1} \dot{\lambda}_t (n+1)^{\delta_t}}{n! \tau(\lambda_t, \delta_t)} + \frac{\lambda_t^n (n+1)^{\delta_t} \dot{\delta}_t \ln(n+1)}{n! \tau(\lambda_t, \delta_t)} - \frac{\lambda_t^n (n+1)^{\delta_t} \dot{\tau}(\lambda_t, \delta_t)}{n! \tau^2(\lambda_t, \delta_t)} \\
 &= \left[n \frac{\dot{\lambda}_t}{\lambda_t} + \dot{\delta}_t \ln(n+1) - \frac{\dot{\tau}(\lambda_t, \delta_t)}{\tau(\lambda_t, \delta_t)} \right] p_n(t) \\
 &= \left[n \frac{\dot{\lambda}_t}{\lambda_t} + \dot{\delta}_t \ln(n+1) - \dot{\mu}_t \right] p_n(t).
 \end{aligned} \tag{5.7}$$

Agora, para aplicarmos a mesma técnica para a dedução com procedimento análogos ao Capítulo 3, com base na definição de derivada

$$\dot{p}_n(t) = \lim_{\Delta t \rightarrow \infty} \frac{p_n(t + \Delta t) - p_n(t)}{\Delta t},$$

e considere incrementos dicotômicos tais que

$$p_n(t + \Delta t) = \pi_{n-1}(t)p_{n-1}(t) + [1 - \pi_n(t)]p_n(t),$$

em que

$$\pi_{n-1}(t) = P[N(t, \Delta t) = 1 | N(t) = n - 1] \quad (5.8)$$

denota a probabilidade de transição de $n - 1$ para n do instante t para $t + \Delta t$. Seguindo essa estrutura, podemos escrever

$$\begin{aligned} p_n(t + \Delta t) - p_n(t) &= \pi_{n-1}(t)p_{n-1}(t) - \pi_n(t)p_n(t) \\ &= \left\{ \pi_{n-1}(t) \left(\frac{n}{n+1} \right)^{\delta_t} \frac{n}{\lambda_t} - \pi_n(t) \right\} p_n(t). \end{aligned}$$

Assim, para Δt suficientemente pequeno, utilizando (5.7), consideramos a aproximação

$$\left\{ \pi_{n-1}(t) \left(\frac{n}{n+1} \right)^{\delta_t} \frac{n}{\lambda_t} - \pi_n(t) \right\} \frac{1}{\Delta t} \approx n \frac{\dot{\lambda}_t}{\lambda_t} + \dot{\delta}_t \ln(n+1) - \dot{\mu}_t.$$

Por conseguinte, concluímos que o processo de Touchard (PrT) pode ser gerado por um incremento Bernoulli $N(t, \Delta t)$ condicionado à contagem $N(t) = n$, cuja probabilidade de sucesso na forma recursiva é expressa como

$$\pi_n(t) \approx \pi_{n-1}(t) \left(\frac{n}{n+1} \right)^{\delta_t} \frac{n}{\lambda_t} + \left\{ \dot{\mu}_t - n \frac{\dot{\lambda}_t}{\lambda_t} - \dot{\delta}_t \ln(n+1) \right\} \cdot \Delta t. \quad (5.9)$$

Para $n = 0$, temos

$$\begin{aligned} \dot{p}_0(t) &= \frac{d}{dt} (p_0(t)) = \frac{d}{dt} \left(\frac{1}{\tau(\lambda_t, \delta_t)} \right) = \frac{d}{dt} (\tau(\lambda_t, \delta_t)^{-1}) \\ &= -\frac{\dot{\tau}(\lambda_t, \delta_t)}{\tau(\lambda_t, \delta_t)} \cdot \frac{1}{\tau(\lambda_t, \delta_t)} \\ &= -\dot{\mu}_t p_0(t), \end{aligned}$$

de modo que a forma recursiva (5.9) se inicia com

$$\pi_0(t) = \dot{\mu}_t \Delta t. \tag{5.10}$$

Observe que se $\lambda_t = \lambda t$ e $\delta_t = 0$, encontramos o processo de Poisson homogêneo $\pi_n(t) = \pi_{n-1}(t)\lambda \cdot \Delta t$, remetendo a incrementos estacionários e independentes.

Porém, o processo de Touchard é gerado por incrementos não-estacionários e dependentes. Consequentemente, o processo gerador de um processo de Poisson ponderado proposto por Balakrishnan e Kozubowski (2008) não se aplica nesta situação.

De fato, considere um intervalo de tempo $[0, t]$ com $t \geq 0$ e uma contagem de Touchard $N = N(1)$. Tomando incrementos independentes e estacionários, de modo que a contagem até o tempo t , $N(t)$, seja proporcional ao comprimento do intervalo $[0, t]$, temos que $N(t)|N = n$ deve seguir uma distribuição Binomial (Balakrishnan e Kozubowski, 2008), ou seja,

$$P[N(t) = n|N] = \binom{N}{n} t^n (1-t)^{N-n},$$

com $0 \leq n \leq N$. Logo, a distribuição do processo $N(t)$ é dada por

$$\begin{aligned} P[N(t) = n] &= \sum_{N \geq n} P[N(t) = n|N] \cdot \frac{\lambda^N (N+1)^\delta}{N! \tau(\lambda, \delta)} \\ &= \sum_{N \geq n} \binom{N}{n} t^n (1-t)^{N-n} \cdot \frac{\lambda^N (N+1)^\delta}{N! \tau(\lambda, \delta)} \end{aligned}$$

$$\begin{aligned}
&= \sum_{N-n \geq 0} \frac{N!}{n!(N-n)!} t^n (1-t)^{N-n} \cdot \frac{\lambda^N (N+1)^\delta}{N! \tau(\lambda, \delta)} \\
&= \sum_{k=0}^{\infty} \frac{1}{n! k!} t^n (1-t)^k \cdot \frac{\lambda^{n+k} (n+k+1)^\delta}{\tau(\lambda, \delta)} \\
&= \frac{(\lambda t)^n}{n!} \sum_{k=0}^{\infty} \frac{[\lambda(1-t)]^k}{k!} \cdot \frac{(n+k+1)^\delta}{\tau(\lambda, \delta)} \\
&= \frac{(\lambda t)^n}{n!} \cdot \frac{T(\lambda(1-t), \delta, n)}{T(\lambda, \delta, 0)},
\end{aligned}$$

em que

$$T(\lambda, \delta, n) = \sum_{k=0}^{\infty} \frac{\lambda^k (n+k+1)^\delta}{k!}. \quad (5.10a)$$

Observação: Na equação (5.10a), se os incrementos forem estacionários e independentes, a contagem $N(t)$ não seguiria uma distribuição de Touchard, exceto para o caso $\delta = 0$.

5.2 A taxa do PrT

Considerando a equação (5.3), a taxa do processo é dada por

$$\begin{aligned}
\frac{E[N(t)]}{t} &= \frac{1}{t} \sum_{n=0}^{\infty} n p_n(t) = \frac{1}{t} \sum_{n=0}^{\infty} \frac{n \lambda_t^n (n+1)^{\delta_t}}{n! \tau(\lambda_t, \delta_t)} \\
&= \frac{1}{t} \sum_{n=0}^{\infty} \frac{\lambda_t^n (n+1)^{\delta_t} (n+1-1)}{n! \tau(\lambda_t, \delta_t)} \\
&= \frac{1}{t} \sum_{n=0}^{\infty} \frac{\lambda_t^n (n+1)^{\delta_t+1}}{n! \tau(\lambda_t, \delta_t)} - \frac{1}{t} \sum_{n=0}^{\infty} \frac{\lambda_t^n (n+1)^{\delta_t}}{n! \tau(\lambda_t, \delta_t)} \\
&= \frac{1}{t} \left(\frac{\tau(\lambda_t, \delta_t + 1)}{\tau(\lambda_t, \delta_t)} - 1 \right) \\
&= \frac{1}{t} \left(\frac{\tau(\lambda_t, \delta_t + 1) - \tau(\lambda_t, \delta_t)}{\tau(\lambda_t, \delta_t)} \right) \\
&= \frac{1}{\tau(\lambda_t, \delta_t)} \left(\frac{\tau(\lambda_t, \delta_t + 1) - \tau(\lambda_t, \delta_t)}{t} \right), \quad (5.11)
\end{aligned}$$

ou

$$\begin{aligned}
\frac{\mathbb{E}[N(t)]}{t} &= \frac{1}{t} \sum_{n=0}^{\infty} n p_n(t) = \frac{1}{t} \sum_{n=0}^{\infty} \frac{n \lambda_t^n (n+1)^{\delta_t}}{n! \tau(\lambda_t, \delta_t)} \\
&= \frac{1}{t} \sum_{n=1}^{\infty} \frac{\lambda_t^n (n+1)^{\delta_t}}{(n-1)! \tau(\lambda_t, \delta_t)} \\
&= \frac{1}{t} \sum_{n'=0}^{\infty} \frac{\lambda_t^{n'+1} (n'+2)^{\delta_t}}{n'! \tau(\lambda_t, \delta_t)} \\
&= \frac{\lambda_t}{t} \sum_{n'=0}^{\infty} \frac{(n'+2)^{\delta_t}}{(n'+1)^{\delta_t}} \cdot \frac{\lambda_t^{n'} (n'+1)^{\delta_t}}{n'! \tau(\lambda_t, \delta_t)} \\
&= \frac{\lambda_t}{t} \sum_{n'=0}^{\infty} \left(\frac{n'+2}{n'+1} \right)^{\delta_t} \cdot p_{n'}(t) \\
&= \frac{\lambda_t}{t} \mathbb{E} \left[\left(\frac{N(t)+2}{N(t)+1} \right)^{\delta_t} \right]. \tag{5.12}
\end{aligned}$$

Assim, para t suficientemente grande, se $\lambda_t \approx \lambda t$ e $\delta_t \rightarrow 0$, ou se δ_t for constante, temos $\frac{\mathbb{E}[N(t)]}{t} \approx \lambda$.

Nessa situação, λ representa uma taxa.

5.3 Momentos

O r -ésimo momento da variável aleatória de Touchard $N(t)$ é uma mistura de Touchard dado por (Matsushita et al., 2019).

$$\begin{aligned}
\mu_r = \mathbb{E}[N^r(t)] &= \sum_{n=0}^{\infty} \frac{(\lambda_t)^n (n+1)^{\delta_t} n^r}{n! \tau(\lambda_t, \delta_t)} = \sum_{n=0}^{\infty} \frac{(\lambda_t)^n (n+1)^{\delta_t} (n+1-1)^r}{n! \tau(\lambda_t, \delta_t)} \\
&= \sum_{n=0}^{\infty} \sum_{j=0}^r \frac{(\lambda_t)^n (n+1)^{\delta_t} \binom{r}{j} (n+1)^j (-1)^{r-j}}{n! \tau(\lambda_t, \delta_t)} \\
&= \sum_{n=0}^{\infty} \sum_{j=0}^r \binom{r}{j} (-1)^{r-j} \cdot \frac{(\lambda_t)^n (n+1)^{\delta_t+j}}{n! \tau(\lambda_t, \delta_t)} \\
&= \sum_{j=0}^r \binom{r}{j} (-1)^{r-j} \cdot \frac{\tau(\lambda_t, \delta_t + j)}{\tau(\lambda_t, \delta_t)}, \tag{5.13}
\end{aligned}$$

e a sua função geradora de momentos é

$$\begin{aligned}
 M_N(t)(q) &= \mathbb{E}[e^{qN(t)}] = \sum_{n=0}^{\infty} \frac{e^{qn} (\lambda_t)^n (n+1)_{\delta_t}}{n! \tau(\lambda_t, \delta_t)} \\
 &= \frac{1}{\tau(\lambda_t, \delta_t)} \sum_{n=0}^{\infty} \frac{(\lambda_t e^q)^n (n+1)_{\delta_t}}{n!} \\
 &= \frac{\tau(\lambda_t e^q, \delta_t)}{\tau(\lambda_t, \delta_t)}. \tag{5.14}
 \end{aligned}$$

$$M_N(t)(q) = \sum_{r=1}^{\infty} \frac{\mu_r q^r}{r!} = \frac{\tau(\lambda_t e^q, \delta_t)}{\tau(\lambda_t, \delta_t)}, \tag{5.15}$$

em que $q \in \mathbb{R}$. Portanto, o processo de Touchard pertence à família de distribuições de séries de potência de dois parâmetros, a média de $N(t)$ pode ser expressa por

$$\mu_t = \mathbb{E}[N(t)] = \frac{\tau(\lambda_t, \delta_t + 1)}{\tau(\lambda_t, \delta_t)} - 1 = \lambda_t \cdot \mathbb{E} \left[\left(\frac{N(t) + 2}{N(t) + 1} \right)^{\delta_t} \right]. \tag{5.16}$$

Variância

Considerando que $\mu_t = \frac{\tau(\lambda_t, \delta_t + 1)}{\tau(\lambda_t, \delta_t)} - 1$, então a variância de $N(t)$ é:

$$\text{Var}[N(t)] = \mathbb{E}[N^2(t)] - [\mathbb{E}(N(t))]^2,$$

mas antes devemos calcular $\mathbb{E}[N^2(t)]$

$$\begin{aligned}
 \mathbb{E}[N^2(t)] &= \binom{2}{0} (-1)^2 \frac{\tau(\lambda_t, \delta_t)}{\tau(\lambda_t, \delta_t)} + \binom{2}{1} (-1) \frac{\tau(\lambda_t, \delta_t + 1)}{\tau(\lambda_t, \delta_t)} + \binom{2}{2} (-1)^0 \frac{\tau(\lambda_t, \delta_t + 2)}{\tau(\lambda_t, \delta_t)} \\
 &= 1 - 2 \frac{\tau(\lambda_t, \delta_t + 1)}{\tau(\lambda_t, \delta_t)} + \frac{\tau(\lambda_t, \delta_t + 2)}{\tau(\lambda_t, \delta_t)}, \tag{5.17}
 \end{aligned}$$

então a variância é dada por

$$\begin{aligned}
 \text{Var}[N(t)] &= E(N^2) - [E(N)]^2 \\
 &= 1 - 2 \frac{\tau(\lambda_t, \delta_t + 1)}{\tau(\lambda_t, \delta_t)} + \frac{\tau(\lambda_t, \delta_t + 2)}{\tau(\lambda_t, \delta_t)} - \mu_t^2 \\
 &= 1 - 2(\mu_t + 1) + \frac{\tau(\lambda_t, \delta_t + 2)}{\tau(\lambda_t, \delta_t)} - \mu_t^2 \\
 &= \frac{\tau(\lambda_t, \delta_t + 2)}{\tau(\lambda_t, \delta_t)} - (\mu_t + 1)^2 \tag{5.18}
 \end{aligned}$$

$$= \frac{\tau(\lambda_t, \delta_t + 2)}{\tau(\lambda_t, \delta_t)} - \left[\frac{\tau(\lambda_t, \delta_t + 1)}{\tau(\lambda_t, \delta_t)} \right]^2. \tag{5.19}$$

Ou ainda, podemos expressar em termos de μ_t como

$$\begin{aligned}
 \text{Var}[N(t)] &= \frac{\tau(\lambda_t, \delta_t + 2)}{\tau(\lambda_t, \delta_t)} - (\mu_t + 1)^2 \\
 &= \frac{\tau(\lambda_t, \delta_t + 2)\tau(\lambda_t, \delta_t + 1)}{\tau(\lambda_t, \delta_t)\tau(\lambda_t, \delta_t + 1)} - (\mu_t + 1)^2 \\
 &= \frac{\tau(\lambda_t, \delta_t + 2)}{\tau(\lambda_t, \delta_t + 1)} \cdot (\mu_t + 1) - (\mu_t + 1)^2 \\
 &= (\mu_t + 1) \cdot \left[\frac{\tau(\lambda_t, \delta_t + 2)}{\tau(\lambda_t, \delta_t + 1)} - (\mu_t + 1) \right]. \tag{5.20}
 \end{aligned}$$

5.4 Fórmula recursiva para a função de probabilidade

Recursivamente, para $n \geq 0$, a probabilidade (5.1) pode ser escrita como

$$\begin{aligned}
 p_{n+1}(t) = P[N(t) = n + 1] &= \frac{(\lambda_t)^{n+1} (n + 2)^{\delta_t}}{(n + 1)! \tau(\lambda_t, \delta_t)} \\
 &= \frac{(\lambda_t)(n + 2)^{\delta_t}}{(n + 1)(n + 1)^{\delta_t}} \cdot \frac{(\lambda_t)^n (n + 1)^{\delta_t}}{n! \tau(\lambda_t, \delta_t)} \\
 &= \frac{(\lambda_t)(n + 2)^{\delta_t}}{(n + 1)(n + 1)^{\delta_t}} \cdot p_n(t) \\
 &= \frac{\lambda_t}{n + 1} \left(\frac{n + 2}{n + 1} \right)^{\delta_t} p_n(t). \tag{5.21}
 \end{aligned}$$

O processo de Touchard é capaz de descrever uma distribuição com contagens inflacionadas a zero. Por exemplo, a partir de (5.21), temos

$$p_1(t) = \lambda_t 2^{\delta_t} p_0(t), \quad (5.22)$$

e

$$p_2(t) = \frac{\lambda_t}{2} \left(\frac{3}{2}\right)^{\delta_t} p_1(t). \quad (5.23)$$

Seja $p_1(t) < p_0(t)$ e $p_1(t) < p_2(t)$, ou seja,

$$\frac{p_1(t)}{p_0(t)} = \frac{\lambda_t^{\delta_t}}{2} < 1, \quad (5.24)$$

e

$$\frac{p_2(t)}{p_1(t)} = \frac{\lambda_t}{2} \left(\frac{3}{2}\right)^{\delta_t} > 1. \quad (5.25)$$

De uma forma mais geral, para algum n , o PrT é inflacionado de zero se $p_n(t) < p_0(t)$ e $p_n(t) < p_{n+1}(t)$, ou seja,

$$\frac{p_n(t)}{p_0(t)} = \frac{(\lambda_t)^n (n+1)^{\delta_t}}{n!} < 1, \quad (5.26)$$

e

$$\frac{p_{n+1}(t)}{p_n(t)} = \frac{\lambda_t}{n+1} \left(\frac{n+2}{n+1}\right)^{\delta_t} > 1. \quad (5.27)$$

5.5 Aspectos computacionais e simulação

Algumas funções do software R para simulação e estimação dos parâmetros do processo de Touchard se encontram disponíveis em <https://1drv.ms/f/s!ApX60k7TMXzeiZha0HEbtX2FDepNLg?e=c3SrHR>. Quanto aos aspectos computacionais referentes à estimação dos parâmetros pelo método

da máxima verossimilhança, discussões detalhadas podem ser encontradas em Matsushita et al. (2019) e Andrade et al. (2021).

Para dar início à simulação para determinada unidade de tempo t e intervalo de tempo Δt , parte-se de uma probabilidade inicial $\pi_0(t)$ definida na equação (5.10). Por construção, como essa probabilidade depende de uma taxa conforme a equação (5.5) estritamente positiva, Δt deve ser suficientemente pequeno para que $\pi_0(t)$ seja, de fato, uma probabilidade.

O problema, porém, é que essa escolha inicial para Δt deve garantir também que $\pi_n(t)$ na equação (5.9) também seja uma probabilidade. Como o segundo termo dessa expressão pode ser negativo, o valor de Δt deve ser tal que a expressão resultante seja sempre positiva.

No entanto, valores de Δt muito próximos de zero geram problemas numéricos que inviabilizam essa abordagem, deteriorando rapidamente a forma recursiva na equação (5.9). Por exemplo, suponha que $t = 1$, $\Delta t = 0.001$, e considere $\lambda_t = \lambda t$ e $\delta_t = \delta$ (invariante no tempo). Nesse caso, observe que o segundo termo na equação (5.9) dependerá da razão $1/t$. Mas na simulação com $\Delta t = 0.001$, o intervalo $(0, t]$ é dividido em 1000 partes, de modo que a simulação consistirá de uma agregação de incrementos $N(t_1, \Delta t), N(t_2, \Delta t), N(t_3, \Delta t), \dots, N(t_{1000}, \Delta t)$, em que $t_j = 0.001j$, para $j = 1, \dots, 1000$. Assim, no início da simulação, o segundo termo de (5.9), que dependeria da razão $1/\Delta t$, quando multiplicado por Δt , poderia retornar valores negativos ou maiores do que 1.

Para solucionar essa questão adotou-se a seguinte estratégia nesse mesmo exemplo. Em vez de iniciarmos a simulação em t_1 , pode-se iniciar em algum instante t_M , simulando diretamente uma Touchard $N(t_M)$ com parâmetros λt_M e δ pelo método da transformação inversa. E, a partir dessa contagem até o instante t_M , aplicamos a forma (5.9) até obtermos a contagem final $N(t)$.

A Figura 5.1 mostra o resultado de uma simulação de uma amostra de tamanho 400 de um processo de Touchard com parâmetros $\lambda_t = 4t$, $\delta_t = 2$. Agrupando as frequências esperadas baixas, o teste χ^2 (com o nível de significância $\alpha = 0.05\%$) de aderência à distribuição de Touchard apresentou p-valor igual a 0.416. Deste modo, não há evidências suficientes para rejeitar a hipótese nula de que os dados seguem uma distribuição de Touchard.

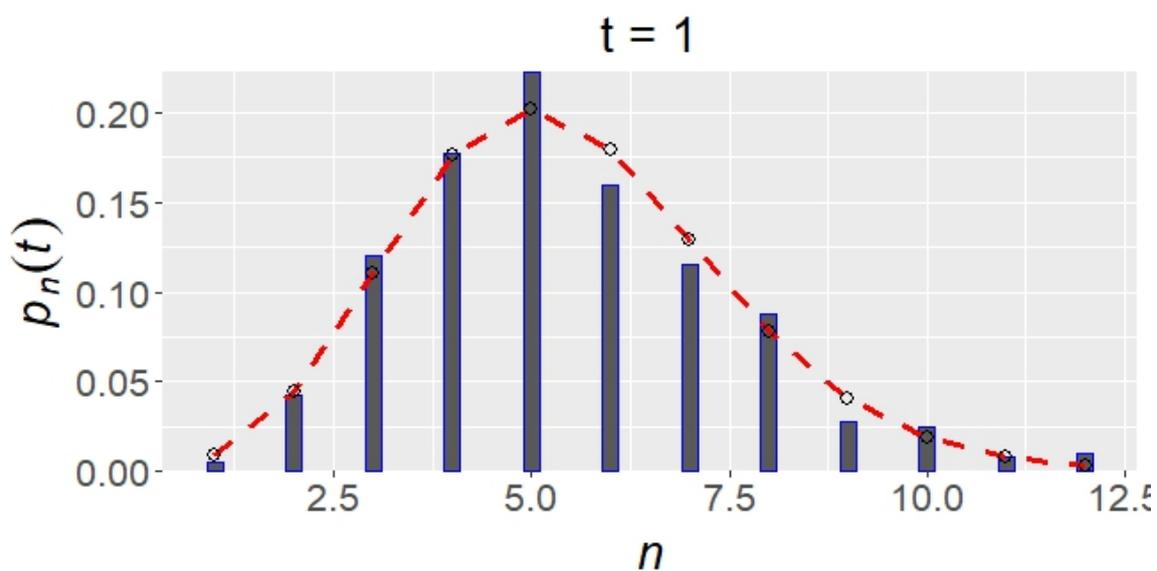


Figura 5.1: Simulação de um processo de Touchard com $\lambda_t = 4t$, $\delta_t = 2$.

Capítulo 6

Aplicação

Neste Capítulo, aplicaremos o modelo estocástico desenvolvido aos dados de acidentes envolvendo veículos motorizados, coletados no Estado de Nova Iorque entre os anos de 2018 e 2022. A análise focará nos acidentes registrados diariamente nos condados de Queens e Kings, permitindo uma avaliação detalhada do comportamento desses eventos em regiões com alta densidade populacional e tráfego intenso. Para cada condado, serão obtidas as estimativas de máxima verossimilhança para o parâmetro do modelo, com o intuito de ajustar adequadamente os dados observados. Além disso, faremos uma comparação entre o modelo de Touchard e o modelo de Poisson, utilizando o Critério de Informação Bayesiano (BIC) como medida de adequação em ambos os condados. Essa análise permitirá verificar a flexibilidade e precisão do modelo de Touchard em relação ao modelo de Poisson tradicional na modelagem de dados de contagem com variabilidade adicional.

6.1 Dados sobre acidentes com veículos motorizados

Os dados sobre acidentes com veículos motorizados foram obtidos no portal de dados abertos do Estado de *New York*, <https://data.ny.gov>. Este conjunto de dados inclui o número de veículos motorizados envolvidos em cada acidente, registrados a partir dos relatórios enviados ao Departamento de Veículos Motorizados de *New York* por motoristas e agências policiais, no período de 2018 a 2022. Cada linha da base de dados representa um evento de colisão, contendo informações detalhadas sobre o acidente. Estão incluídas todas as colisões com veículos motorizados registradas pela polícia no Es-

tado de *New York*. O preenchimento do relatório policial (MV104-AN) é obrigatório para colisões que resultem em ferimentos, mortes, ou quando os danos materiais superam 1000 dólares.

Para nossa análise, consideramos os acidentes registrados diariamente nos condados de *Queens* e *Kings*. No condado de *Queens*, trabalhamos com janelas de tempo variando de 5 a 60 minutos, com incrementos de 5 minutos, a partir das 18:00 horas, utilizando uma amostra de 1826 dias. Para o condado de *Kings*, a análise foi realizada da mesma forma, mas a partir das 17:00 horas, com a mesma dimensão amostral.

Analisando a frequência de acidentes por condado, observamos que *Queens* registrou o maior número de acidentes com veículos motorizados, seguido por *Kings*, no período de 2018 a 2022. Em seguida, realizamos uma análise temporal para identificar a concentração de acidentes ao longo do dia, além de uma análise detalhada dos minutos para entender a distribuição global dos registros, com destaque para concentrações decorrentes de arredondamentos nos horários.

As Figuras 6.1 e 6.2 mostram que o maior número de acidentes ocorre entre 06:00 e 21:00. Quanto aos minutos, há uma maior concentração nos minutos 30 e 60, possivelmente devido ao arredondamento. Além disso, os múltiplos de 5 também apresentam subconcentrações, com destaque especial para os minutos 0, 15, 30 e 45, resultado desse fenômeno de arredondamento natural.

Esse tipo de concentração é comum em outros processos, como a chegada de consumidores após o desembarque de ônibus ou nos horários de abertura e fechamento de estabelecimentos, criando complexidades em processos estocásticos e influenciando as probabilidades dos incrementos.

Assim, o registro de eventos em "pacotes" pode induzir não estacionariedade — já que a ocorrência de sucessos depende do instante do registro — e dependência — pois as ocorrências de sucessos em instantes próximos tendem a serem registrados em um horário aproximado comum.

Por fim, analisamos as contagens de acidentes em horários específicos nos dois condados. Para *Queens*, focamos no intervalo entre 18:00 e 19:00; para *Kings*, entre 17:00 e 18:00. Utilizamos estimativas de máxima verossimilhança e intervalos de confiança para essa análise.

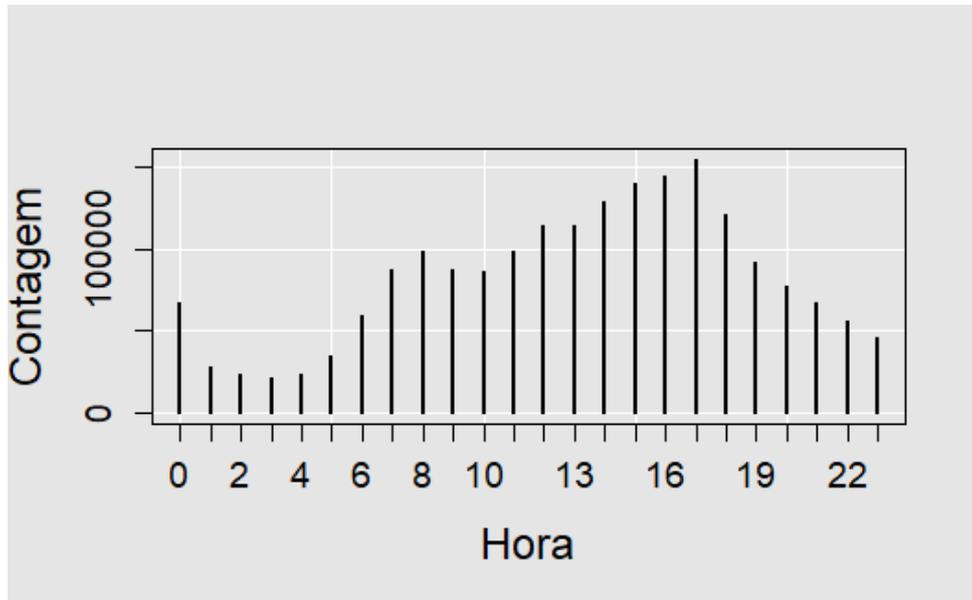


Figura 6.1: Distribuição horário global (em horas).

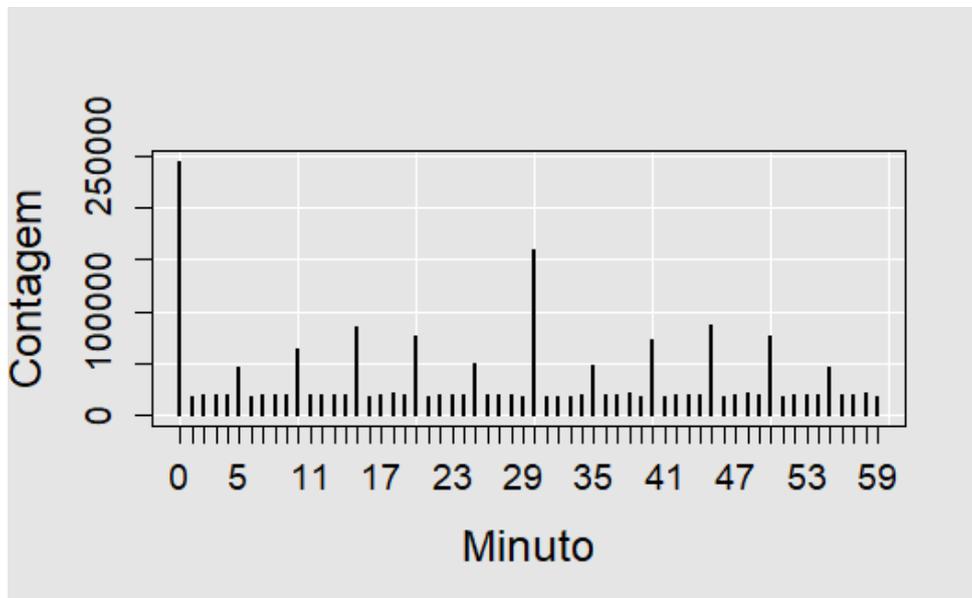


Figura 6.2: Distribuição global por minuto.

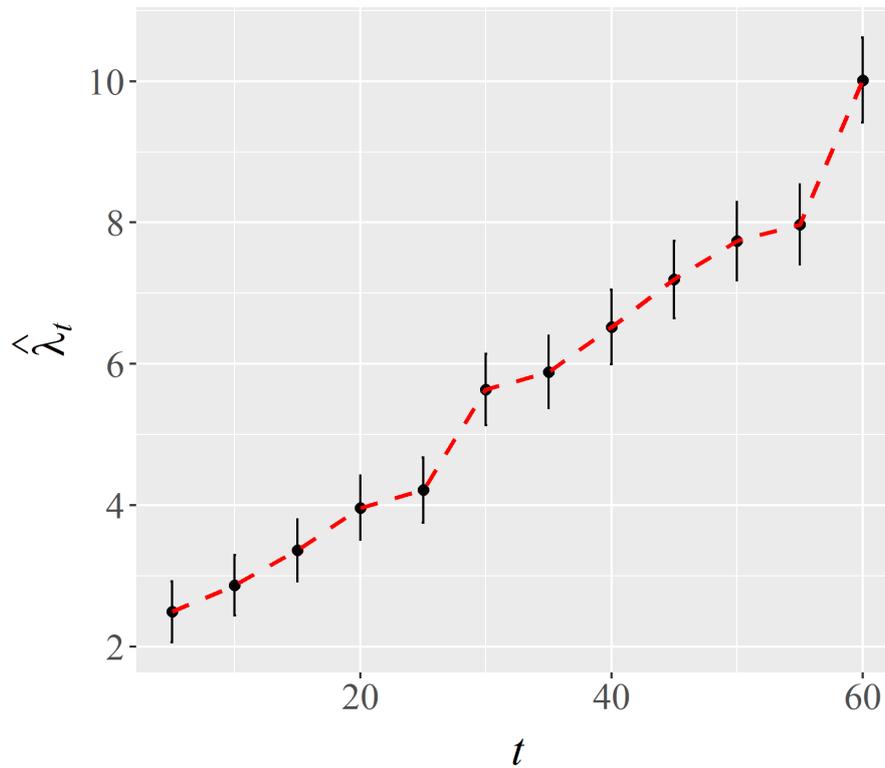


Figura 6.3: Estimativas de máxima verossimilhança de λ_t , com $t = 5$ a 60 minutos respectivamente, com incremento de 5 minutos a partir das 18:00. As barras de erro referem-se aos intervalos de confiança de 99,7% com base na propriedade de normalidade assintótica.

6.2 Observados no Condado de Queens

As estimativas de máxima verossimilhança para o parâmetro λ_t do modelo de Touchard, com t variando de 5 a 60 minutos em incrementos de 5 minutos a partir das 18:00, mostraram-se altamente significativas. Utilizando a normalidade assintótica dos estimadores, verificamos que os p -valores são inferiores a 10^{-7} , e as barras de erro representam intervalos de confiança de 99,7%.

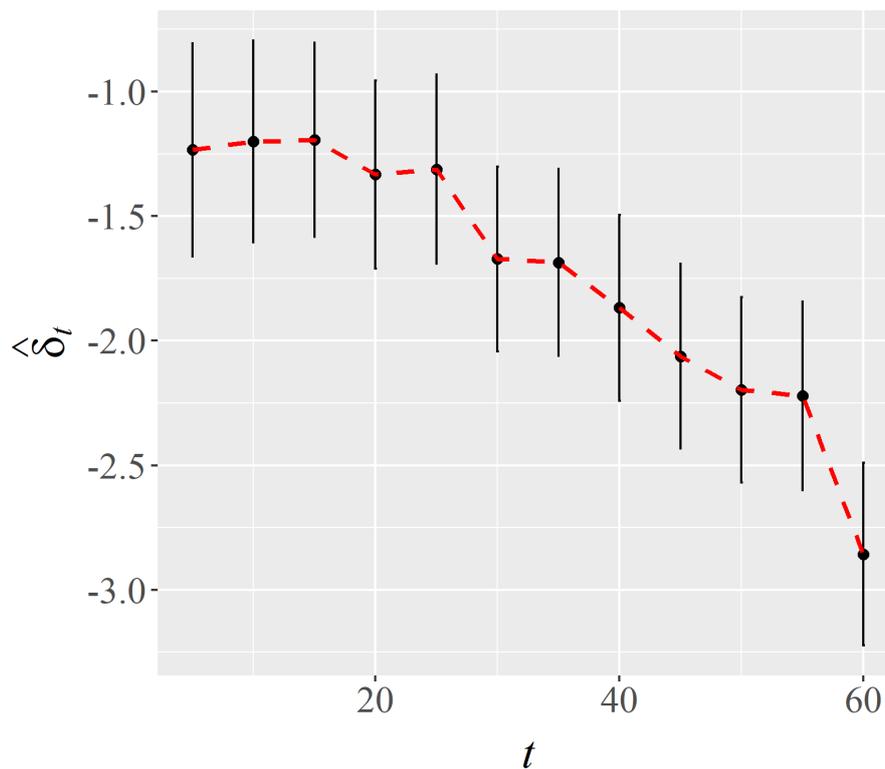


Figura 6.4: Estimativas de máxima verossimilhança de δ_t , com $t = 5$ a 60, com incremento de 5 minutos a partir das 18:00. As barras de erro referem-se aos intervalos de confiança de 99,7% com base na propriedade de normalidade assintótica.

As estimativas de máxima verossimilhança para o parâmetro δ_t do modelo de Touchard também se mostraram altamente significativas, com t variando de 5 a 60 minutos e p -valores inferiores a 10^{-7} . As barras de erro refletem intervalos de confiança de 99,7%.

A Figura 6.5 ilustra as variâncias do número de acidentes observados no condado de *Queens* em diferentes janelas de tempo.

Das 18:00 até 18 : 00 + t , durante o período de 2018 a 2022. As bolinhas sólidas pretas representam

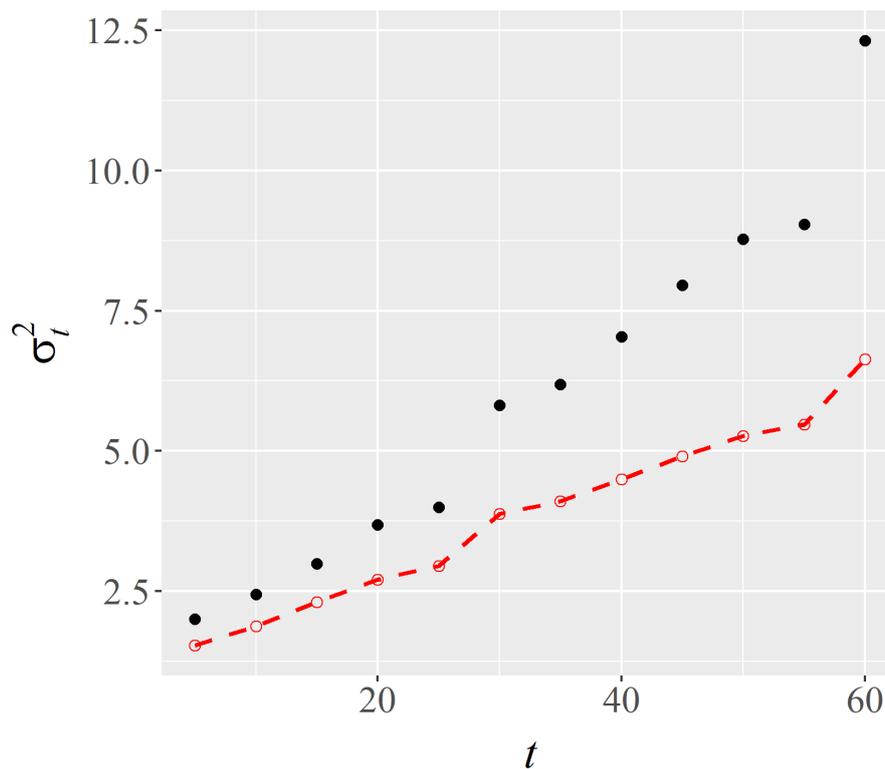


Figura 6.5: Variâncias do número de acidentes observados no condado de *Queens* (bolas sólidas), das 18:00 às 18:(00 + t), onde $t = 5$ a 60 minutos respectivamente, por incremento de 5 minutos. As bolas vazias denotam a média amostral correspondente, e a linha tracejada descreve sua evolução ao longo do tempo t , evidenciando a superdispersão em comparação com o processo de Poisson.

as variâncias observadas, enquanto as bolinhas vazias indicam a média amostral correspondente. A linha tracejada vermelha mostra a evolução da média ao longo do tempo t , evidenciando uma superdispersão em relação ao processo de Poisson.

Para demonstrar a eficiência e simplicidade do modelo, a Figura 6.6 apresenta as distribuições empíricas diárias de acidentes (barras) comparadas com as contagens esperadas pelo modelo de Touchard (\circ). O gráfico evidencia o bom ajuste proporcionado pela flexibilidade do processo de Touchard em relação ao modelo de Poisson. As barras representam as frequências relativas diárias de acidentes registrados no condado de *Queens*, enquanto as linhas pontilhadas destacam a forma da distribuição do modelo de Touchard.

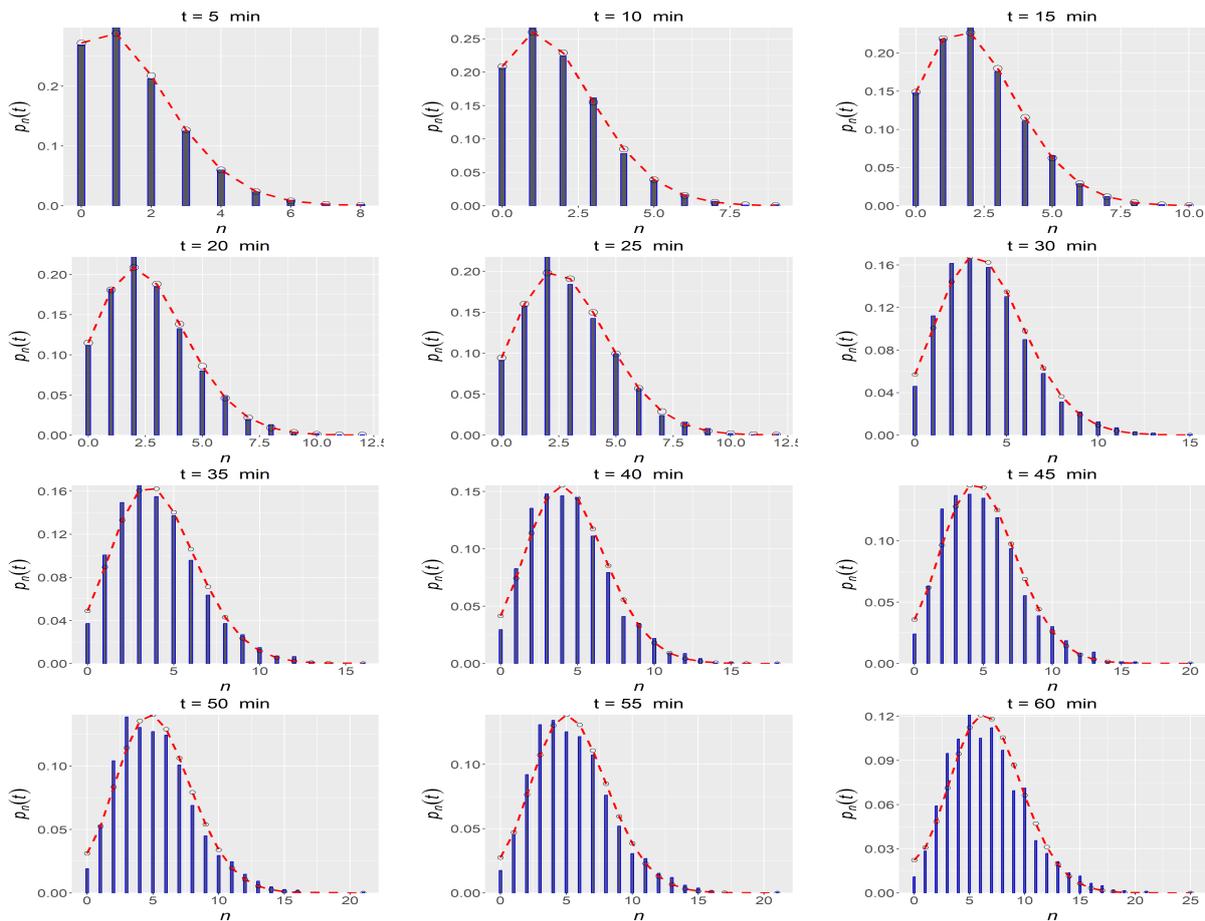


Figura 6.6: Barras verticais: Frequências relativas diárias de acidentes registrados no condado de *Queens*, das 18:00 às 18:(00 + t), em que $t = 5, 10, 15, \dots, 60$ minutos respectivamente. As linhas pontilhadas destacam a forma da distribuição de Touchard (\circ).

Na Tabela 6.1, são apresentados os resultados das estimativas de máxima verossimilhança (EMV) dos parâmetros do processo de Touchard, juntamente com os respectivos erros padrão (s.e.), a média e a variância amostral do número de acidentes observados no condado de *Queens*. Além disso, é feita uma comparação entre os Critérios de Informação Bayesiana (BIC) dos modelos de Touchard e Poisson, e o Teste de Razão de Verossimilhança de Wilk (TRV), considerando H_0 : Poisson vs H_1 : Touchard, com os respectivos p -valores.

Tabela 6.1: Resumo das estimativas de máxima verossimilhança dos parâmetros do processo de Touchard e comparação com o modelo de Poisson no condado de *Queens*.

t	$\hat{\lambda}$	s.e.($\hat{\lambda}$)	$\hat{\delta}$	s.e.($\hat{\delta}$)	$\hat{\mu}_t$	$\hat{\sigma}_t^2$	BIC		TRV statistic	p-value
							Touchard	Poisson		
5	2.49	0.143	-1.23	0.143	1.5	2.00	5957.3	6018.8	69.1	0.000
10	2.87	0.143	-1.20	0.135	1.9	2.44	6410.3	6475.2	72.4	0.000
15	3.36	0.145	-1.19	0.130	2.3	2.99	6878.7	6947.4	76.1	0.000
20	3.96	0.152	-1.33	0.126	2.7	3.68	7288.9	7379.9	98.5	0.000
25	4.21	0.153	-1.31	0.126	3.0	4.00	7475.0	7561.6	94.1	0.000
30	5.63	0.168	-1.67	0.123	3.9	5.82	8201.0	8343.1	149.6	0.000
35	5.88	0.170	-1.69	0.125	4.1	6.19	8325.3	8464.8	147.0	0.000
40	6.52	0.176	-1.87	0.124	4.5	7.03	8574.6	8743.1	176.1	0.000
45	7.19	0.181	-2.06	0.124	4.9	7.96	8818.5	9020.6	209.6	0.000
50	7.74	0.186	-2.20	0.124	5.3	8.78	9008.7	9230.7	229.5	0.000
55	7.97	0.188	-2.22	0.126	5.5	9.04	9075.5	9293.0	225.0	0.000
60	10.02	0.200	-2.86	0.122	6.6	12.31	9673.5	10018.1	352.1	0.000

Se o p -valor for menor ou igual ao nível de significância, rejeite a hipótese nula e conclua que há uma associação estatisticamente significativa entre as variáveis.

A Tabela 6.1 resume as EMV dos parâmetros, destacando que as variâncias amostrais foram consistentemente superiores às médias em todos os intervalos de tempo t , o que indica uma superdispersão em relação à distribuição de Poisson. Ao comparar os modelos de Poisson e Touchard, observou-se que o modelo de Touchard apresenta valores de BIC mais baixos em todos os intervalos, sugerindo um melhor ajuste aos dados. Os resultados do teste de Razão de Verossimilhança também corroboram essa conclusão, rejeitando a hipótese de que as contagens seguem uma distribuição de Poisson com alto nível de significância estatística (valores de p -valor menores que 10^{-7}). Além disso, as estimativas dos parâmetros, baseadas na normalidade assintótica dos estimadores de máxima verossimilhança,

demonstraram-se altamente significativas, fortalecendo a validade do modelo de Touchard para os dados analisados.

6.2.1 Parâmetros Estimados

A estimativa de máxima verossimilhança de λ ($\hat{\lambda}$) representa a taxa média de ocorrência de acidentes, e observa-se que ela aumenta com o tempo. Isso é esperado, pois intervalos de tempo maiores tendem a registrar mais acidentes. O erro padrão de λ (s.e. ($\hat{\lambda}$)) indica a precisão da estimativa. Erros padrão menores sugerem maior confiança nos resultados. O parâmetro δ ($\hat{\delta}$) captura a superdispersão dos dados, ou seja, quando a variabilidade é maior do que a prevista pelo modelo de Poisson. Um valor negativo de δ indica que o modelo de Touchard consegue ajustar adequadamente essa superdispersão. Assim como para λ , o erro padrão de δ (s.e. ($\hat{\delta}$)) reflete a precisão da estimativa, sendo menores os erros padrão, maior a confiança.

6.2.2 Momentos Amostrais

A média observada ($\hat{\mu}_t$) dos acidentes aumenta com o tempo, conforme esperado, refletindo a maior contagem de acidentes em intervalos maiores. A variância da amostra ($\hat{\sigma}_t^2$) também aumenta e é consistentemente maior que a média, confirmando a presença de superdispersão nos dados. Esse comportamento justifica o uso do modelo de Touchard, que lida melhor com a superdispersão do que o modelo de Poisson.

6.2.3 Comparação de Modelos

O BIC é uma métrica de seleção de modelos que avalia a qualidade do ajuste dos modelos, penalizando modelos mais complexos. Para todos os intervalos de tempo, o modelo de Touchard apresenta valores de BIC menores do que o modelo de Poisson, indicando um ajuste superior devido à capacidade do modelo de capturar a superdispersão por meio do parâmetro δ .

O BIC é dado pela fórmula

$$BIC = k \log(n) - 2 \log \mathcal{L}$$

em que,

- k é o número de parâmetros no modelo (1 para Poisson e 2 para Touchard);
- $n = 1826$ é o número de observações;
- $\log \mathcal{L}$ é o log da verossimilhança máxima.

Podemos rearranjar a fórmula do BIC para calcular $\log \mathcal{L}$ (Log-verossimilhança)

$$\log \mathcal{L} = -\frac{BIC - k \log(n)}{2}.$$

Agora vamos calcular as log-verossimilhanças para os dois modelos. Para $t = 5$:

1. Modelo de Touchard ($k = 2$),

$$BIC_{Touchard} = 5957.3,$$

$$\log \mathcal{L}_{Touchard} = -\frac{5957.3 - 2 \log(1826)}{2}.$$

Calculando $\log(1826)$

$$\log(1826) \approx 7.509,$$

$$\log \mathcal{L}_{Touchard} = -\frac{5957.3 - 2 \times 7.509}{2} = -\frac{5957.3 - 15.018}{2} = -\frac{5942.282}{2} = -2971.141.$$

1. Modelo de Poisson ($k = 1$),

$$BIC_{Poisson} = 6018.8,$$

$$\log \mathcal{L}_{Poisson} = -\frac{6018.8 - 1 \log(1826)}{2},$$

$$\log \mathcal{L}_{Poisson} = -\frac{6018.8 - 7.509}{2} = -\frac{6011.291}{2} = -3005.646.$$

6.2.4 A Estatística do Teste de Razão de Verossimilhança (TRV)

O TRV compara os ajustes dos modelos de Poisson e Touchard, com valores mais altos indicando que o modelo de Touchard se ajusta significativamente melhor. Todos os p -valores são extremamente pequenos ($p < 0.001$), permitindo a rejeição da hipótese nula de que os dados seguem uma distribuição de Poisson em favor do modelo de Touchard.

A estatística D do TRV é dada por

$$D = 2 \times (\log \mathcal{L}_{Touchard} - \log \mathcal{L}_{Poisson}).$$

Substituindo os valores calculados

$$D = 2 \times (-2971.141 - (-3005.646)) = 2 \times (34.505) = 69.01.$$

Esse valor segue uma distribuição qui-quadrado (χ^2) com 1 grau de liberdade (diferença no número de parâmetros entre os dois modelos), e o p -valores correspondente é extremamente pequeno, confirmando a superioridade do modelo de Touchard.

O p -valores pode ser calculado como

$$p\text{-value} = P(\chi_1^2 \geq 69.01).$$

6.2.5 Significância dos Parâmetros ($\hat{\lambda}$ e $\hat{\delta}$)

As estimativas de λ e δ são altamente significativas do ponto de vista estatístico. Para $t = 5$, os rácios entre a estimativa e seu erro padrão foram

$$Z_{\hat{\lambda}} = \frac{\hat{\lambda}}{s.e.(\hat{\lambda})} = \frac{2.49}{0.143} = 17.43,$$

$$Z_{\hat{\delta}} = \frac{\hat{\delta}}{s.e.(\hat{\delta})} = \frac{-1.23}{0.143} = 8.63.$$

Esses valores representam a estatística z para cada estimativa, e quando combinados com a suposição de que os estimadores de máxima verossimilhança têm distribuição normal assintótica, podemos verificar

sua significância. Valores de Z superiores a aproximadamente 3.29 correspondem a p -valores menores que 0.001. Aqui, com valores de Z muito maiores (como 17.43 e 8.63), os p -valores serão extremamente pequenos, confirmando que ambas as estimativas são altamente significativas do ponto de vista estatístico (geralmente $p < 10^{-7}$).

6.3 Observados no Condado de Kings

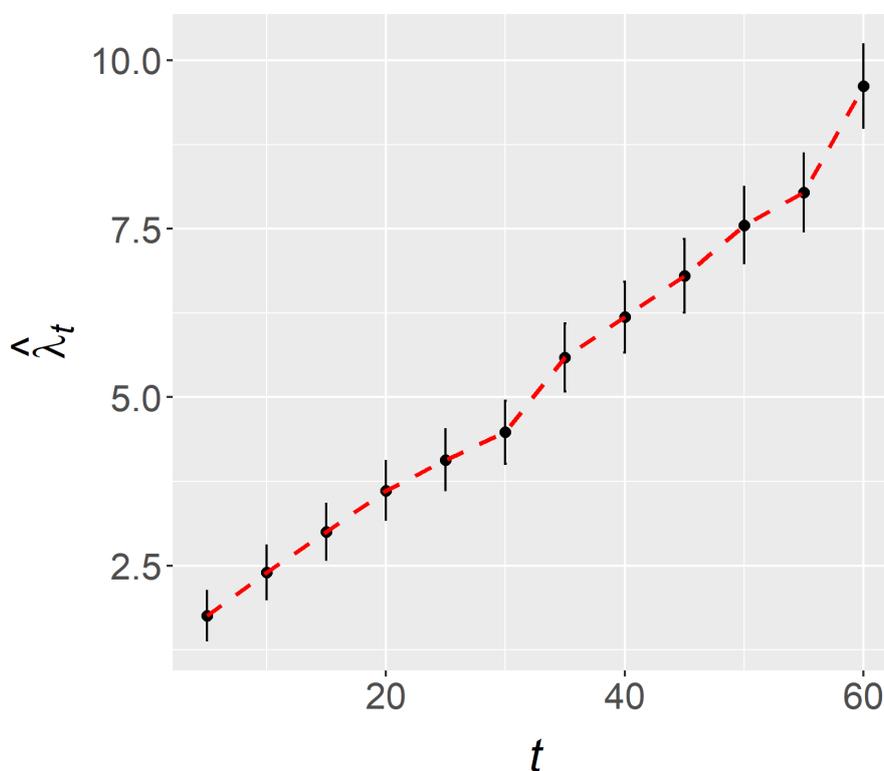


Figura 6.7: Estimativas de máxima verossimilhança de λ_t , com $t = 5$ a 60 minutos respectivamente, com incremento de 5 minutos a partir das 17:00. As barras de erro referem-se aos intervalos de confiança de 99,7% com base na propriedade de normalidade assintótica.

Para o caso do condado de *Kings*, as EMV do parâmetro δ_t no modelo de Touchard, com t variando de 5 a 60 minutos em incrementos de 5 minutos a partir das 17:00, são estatisticamente significativas, com p -valores inferiores a 10^{-7} . As barras de erro representam intervalos de confiança de 99,7%. As EMV para δ_t são altamente significativas, com p -valores abaixo de 10^{-7} , e as barras de erro indicam intervalos de confiança de 99,7%.

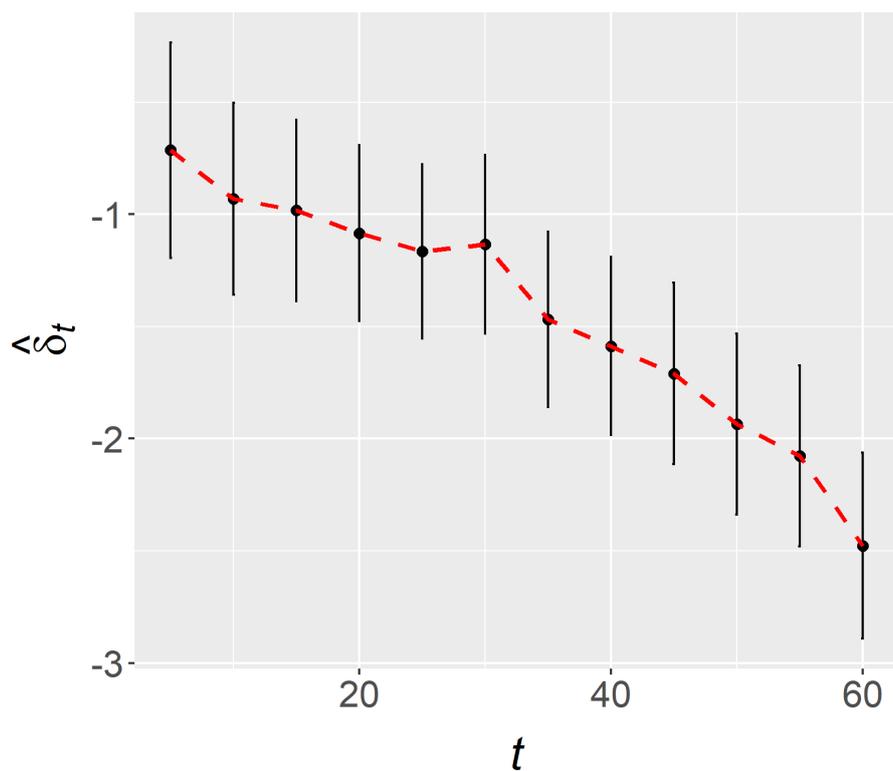


Figura 6.8: Estimativas de máxima verossimilhança de δ_t , com $t = 5$ a 60, com incremento de 5 minutos a partir das 17:00. As barras de erro referem-se aos intervalos de confiança de 99,7% com base na propriedade de normalidade assintótica.

Na sequência, a Figura 6.9 mostra as variâncias do número de acidentes observados no condado de *Kings* em função de t .

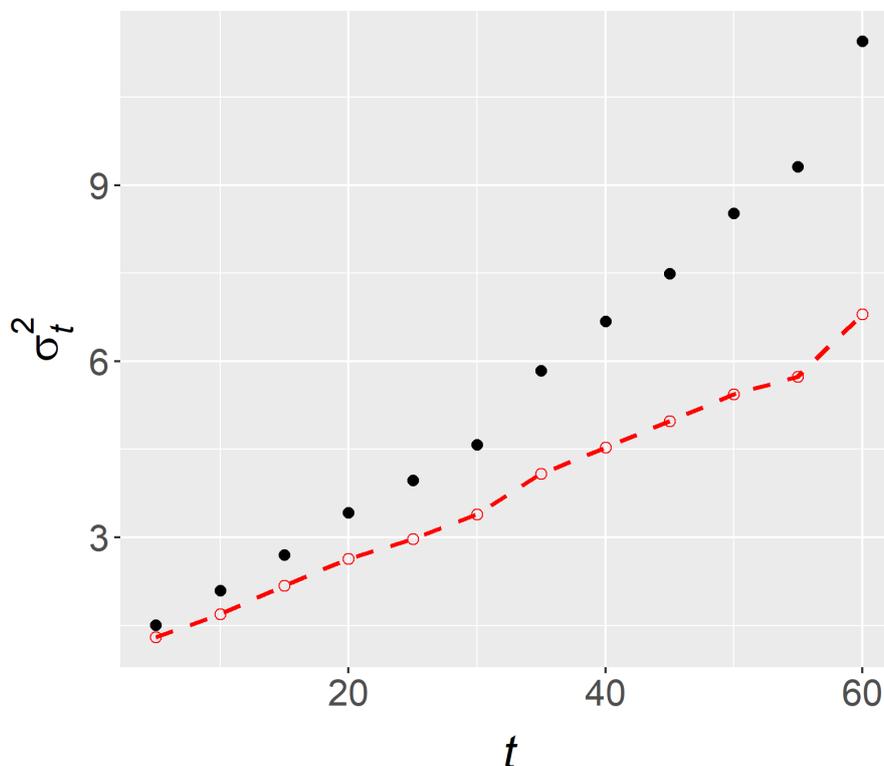


Figura 6.9: Variâncias do número de acidentes observados no condado de *Kings* (bolinhas sólidas), das 17:00 às 18:(00 + t), onde $t = 5$ a 60 minutos respectivamente, por incremento de 5 minutos. As bolinhas vazias denotam a média amostral correspondente, e a linha tracejada descreve sua evolução ao longo do tempo t , evidenciando a superdispersão em comparação com o processo de Poisson..

As variâncias (bolinhas sólidas pretas) correspondem ao período das 17:00 até 17:00 + t , no intervalo de 2018 a 2022, e indicam superdispersão em relação ao processo de Poisson. As bolinhas vazias representam a média amostral, e a linha tracejada avermelhada destaca sua evolução ao longo do tempo, como acontece para o caso de condado de *Queens*.

A fim de ilustrar a performance do modelo, a Figura 6.10 compara as distribuições empíricas diárias de acidentes (barras) com as contagens esperadas segundo o modelo de Touchard (\circ), demonstrando um bom ajuste devido à maior flexibilidade desse modelo em relação ao de Poisson.

Distribuição dos acidentes ao longo do tempo. cada gráfico mostra as frequências relativas de aci-

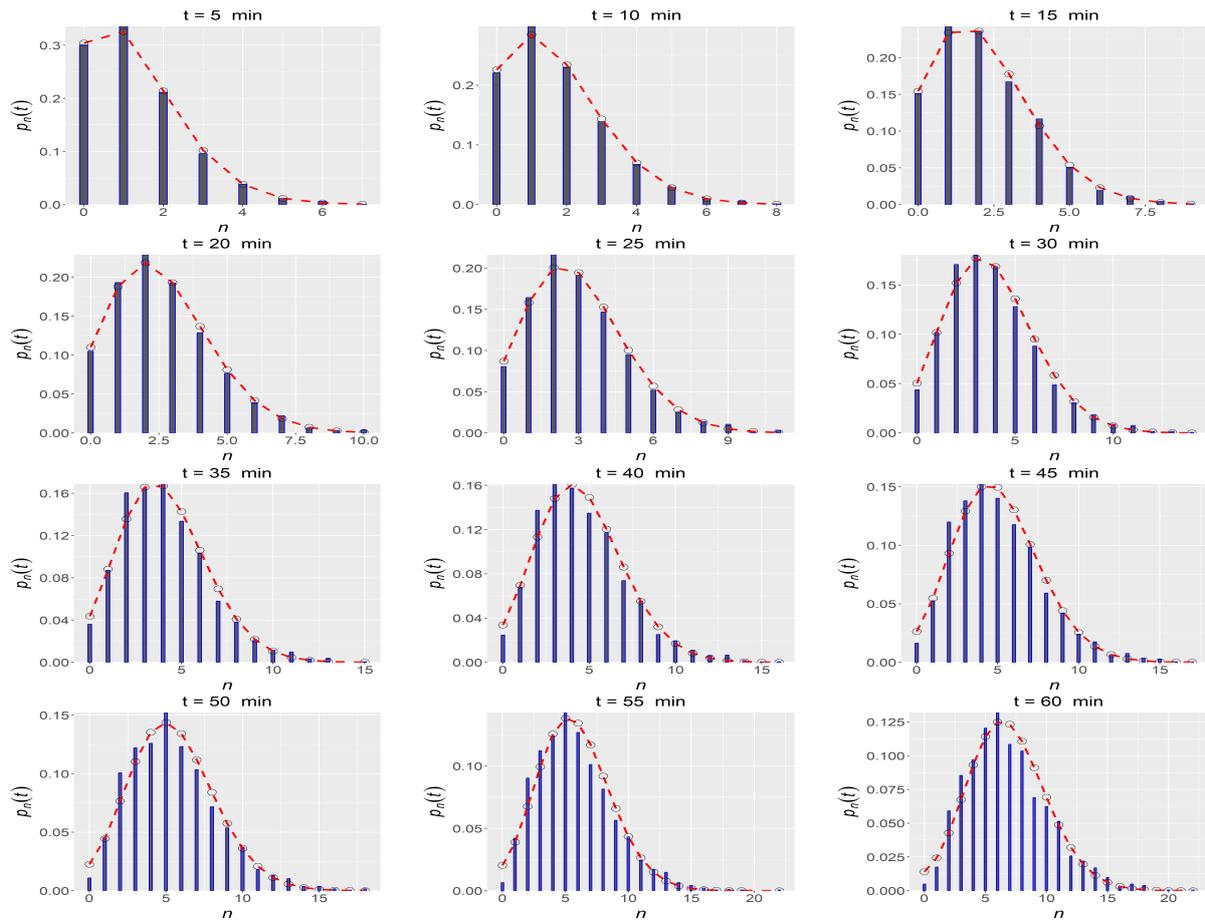


Figura 6.10: Barras verticais: Frequências relativas diárias de acidentes registrados no condado de Kings, das 17:00 às 17:(00 + t), em que $t = 5, 10, 15, \dots, 60$ minutos respectivamente. As linhas pontilhadas destacam a forma da distribuição de Touchard (\circ).

dentos ocorridos em diferentes períodos, variando de 5 minutos até 60 minutos. O comportamento dos acidentes ao longo do tempo é comparado com as previsões do modelo de Touchard, que tenta capturar a tendência dos dados reais. Em geral, os círculos representam a previsão do modelo de Touchard e, ao compará-los com as barras que mostram os dados reais, pode-se avaliar a capacidade do modelo em capturar a distribuição dos acidentes. Se os círculos estiverem próximos das barras, isso indica que o modelo está fazendo boas previsões. À medida que o tempo de intervalo aumenta (por exemplo, de 5 a 60 minutos), a dispersão das frequências de acidentes pode variar. O ajuste do modelo também pode ser diferente conforme o intervalo aumenta. A superdispersão, que o modelo de Touchard tenta capturar, pode ser evidente se houver grandes variações entre as médias e as variâncias das contagens de acidentes. Em suma o modelo de Touchard ajusta bem os dados, especialmente quando as contagens reais (barras) e as contagens esperadas (círculos) estão próximas. Em alguns intervalos, as previsões do modelo podem se afastar ligeiramente das observações, sugerindo a presença de padrões não completamente capturados, como flutuações em horários de pico ou variações sazonais.

Tabela 6.2: Estimativas de máxima verossimilhança dos parâmetros do processo de Touchard e respectivos erros padrão (s.e.), a média amostral e a variância do número de acidentes observados no condado de *Kings*. Comparação dos Critérios de Informação Bayesiana (BIC) sob os modelos Touchard e Poisson, além do teste de Razão de Verossimilhança de Wilk (TRV) (H_0 : Poisson vs H_1 : Touchard), com os respectivos p -valores.

t	$\hat{\lambda}$	s.e.($\hat{\lambda}$)	$\hat{\delta}$	s.e.($\hat{\delta}$)	$\hat{\mu}_t$	$\hat{\sigma}_t^2$	BIC		TRV statistic	p-value
							Touchard	Poisson		
5	1.76	0.123	-0.71	0.160	1.3	1.51	5471.0	5482.6	19.1	0.000
10	2.40	0.133	-0.93	0.143	1.7	2.09	6125.4	6157.8	39.9	0.000
15	3.00	0.139	-0.98	0.134	2.2	2.70	6698.0	6739.8	49.2	0.000
20	3.62	0.146	-1.08	0.131	2.6	3.42	7164.5	7218.8	61.9	0.000
25	4.07	0.151	-1.16	0.129	3.0	3.97	7458.9	7523.1	71.8	0.000
30	5.15	0.163	-1.35	0.131	3.8	5.29	8036.2	8118.9	90.2	0.000
35	5.59	0.168	-1.47	0.130	4.1	5.84	8227.6	8325.3	105.2	0.000
40	6.19	0.175	-1.59	0.132	4.5	6.68	8486.2	8595.2	116.6	0.000
45	6.80	0.182	-1.71	0.135	5.0	7.49	8714.7	8834.0	126.8	0.000
50	7.55	0.189	-1.94	0.135	5.4	8.52	8963.7	9112.9	156.6	0.000
55	8.04	0.193	-2.08	0.134	5.7	9.31	9132.4	9300.4	175.6	0.000
60	9.62	0.207	-2.48	0.138	6.8	11.46	9546.8	9758.1	218.9	0.000

A Tabela 6.2 resume as EMV dos parâmetros, destacando que as variâncias amostrais foram con-

sistentemente superiores às médias em todos os intervalos de tempo t , o que indica uma superdispersão em relação à distribuição de Poisson. Ao comparar os modelos de Poisson e Touchard, observou-se que o modelo de Touchard apresenta valores de BIC mais baixos em todos os intervalos, sugerindo um melhor ajuste aos dados. Os resultados do teste de Razão de Verossimilhança também corroboram essa conclusão, rejeitando a hipótese de que as contagens seguem uma distribuição de Poisson com alto nível de significância estatística (valores de p -valor menores que 10^{-7}). Além disso, as estimativas dos parâmetros, baseadas na normalidade assintótica dos estimadores de máxima verossimilhança, demonstraram-se altamente significativas, fortalecendo a validade do modelo de Touchard para os dados analisados.

Capítulo 7

Conclusão

O processo de Touchard, derivado da modificação de alguns pressupostos do clássico processo de Poisson, mostrou-se uma ferramenta promissora para modelar contagens não-Poisson. A distribuição proposta amplia o conjunto de processos estocásticos de tempo contínuo aplicáveis a eventos pontuais em dados reais. Sua flexibilidade em lidar com dados subdispersos ou superdispersos, bem como com caudas leves ou pesadas e excesso de zeros, torna esse modelo particularmente interessante. Diferentemente do processo de Poisson, no qual o parâmetro λ controla a média da distribuição, o de Touchard introduz um parâmetro adicional, δ , que permite ajustar a dispersão. Como consequência, o modelo de Touchard pode representar dados superdispersos, subdispersos e com excesso de zeros, como evidenciado em nosso estudo de caso, ilustrado nas Figuras 6.6 e 6.10.

Embora a distribuição de Touchard pertença à classe dos processos de Poisson ponderados, ela difere da abordagem proposta por Balakrishnan e Kozubowski (2008), que é gerada por uma taxa proporcional ao tempo com incrementos independentes associados a uma função de peso. Nosso modelo proposto se destaca por lidar com incrementos não estacionários e dependentes, o que o diferencia dos processos estocásticos tradicionais.

Enquanto as probabilidades de transição no processo de Poisson são determinadas por uma equação de diferença-diferencial com taxa constante, no processo de Touchard, essas probabilidades são obtidas de forma recursiva. Esta é uma das principais contribuições desta dissertação, que oferece uma nova abordagem para descrever processos estocásticos com incrementos não estacionários e dependentes, es-

tendendo a classe dos processos de Poisson ponderados.

Nos dados analisados sobre o número de acidentes diários com veículos motorizados em dois condados do estado de *New York*, o BIC indicou que o modelo de Touchard é o mais adequado em todos os intervalos de tempo considerados. Os resultados do teste de razão de verossimilhança também confirmaram esse ajuste, rejeitando a hipótese de que as contagens seguem uma distribuição de Poisson, com resultados estatisticamente significativos. Embora existam outras distribuições que acomodam subdispersão, superdispersão e excesso de zeros, focamos na comparação com a distribuição de Poisson devido à sua ampla utilização na literatura de transporte, especialmente na modelagem de acidentes. Situações com problemas de registros observacionais, como o arredondamento de horários, podem gerar um pico artificial de acidentes, influenciando o processo empírico.

Por fim, destacamos que o processo de Touchard oferece uma maior flexibilidade. Além disso, os cálculos numéricos podem ser facilmente programados em qualquer linguagem. Utilizamos o software R para nosso estudo, cujas funções para simulação e estimação dos parâmetros do processo de Touchard se encontram disponíveis em <https://1drv.ms/f/s!Apx60k7TMXzeiZha0HEbtX2FDepNLg?e=c3SrHR>.

Apêndice A

Regressão de Poisson

A.1 Regressão de Poisson

Em seguida vamos fazer uma pequena introdução de Modelos Lineares Generalizados, (GLM, do inglês Generalized Linear Models). Dando ênfase detalhadamente do modelo Poisson e binomial negativa, os modelos lineares generalizados (MLG) são uma extensão dos modelos de regressão linear clássicos, que foram propostos por Nelder e Wedderburn (1972). Em que permite essencialmente analisar a relação entre um conjunto de variáveis independentes N_1, \dots, N_m (preditor linear) e a variável dependente N_i chamada de variável resposta. A variável resposta N_i deve seguir uma distribuição pertencente à família exponencial.

Família Exponencial

Dizemos que a distribuição de uma variável aleatória N (discreta ou contínua) é membro da família exponencial (FE), se sua função de massa de probabilidade (fmp) ou a sua fdp é dada por

$$f(n; \theta, \phi) = \exp \left[\frac{n\theta - b(\theta)}{a(\phi)} + C(n, \phi) \right], \quad (\text{A.1})$$

sendo θ o parâmetro canônico, tal que $\theta \in \Theta$ (espaço paramétrico), $b(\theta)$ e $C(n, \phi) : \mathbb{R} \rightarrow \mathbb{R}$, A é um conjunto que não depende de θ e ϕ . Por sua vez, $a(\phi)$ é uma função que depende apenas do parâmetro de precisão ϕ ($\phi > 0$). A função $b(\theta)$ é comumente conhecida como geradora de cumulantes e, tem uma especial importância nos modelos lineares generalizados, é uma função diferencial, cujo seu suporte não

deve depender dos parâmetros. Os MLG consistem em três componentes (Nelder e Wedderburn, 1972).

- Um componente aleatório N_i chamado variável resposta, a distribuição de N pertence à família exponencial. Em que a esperança da variável aleatória N é dada da forma

$$E(N_i) = \mu_i = \dot{b}(\theta_i). \quad (\text{A.2})$$

E a variância da variável aleatória N é dada da forma:

$$\text{Var}(N_i) = a(\phi)\ddot{b}(\theta_i) \quad (\text{A.3})$$

. Neste caso, vê-se que a variância de N é o produto entre $\ddot{b}(\theta)$ (que depende apenas do parâmetro canônico θ) e $a(\phi)$ (que depende apenas de ϕ). Observe que, a segunda derivada da função geradora de cumulantes $\ddot{b}(\theta)$ é igualmente conhecido como função de variância (notação, $V(\mu) = \ddot{b}(\theta)$).

- Componente sistemático: Variáveis explicativas na forma de uma soma linear de seus efeitos, que é uma função linear das variáveis preditoras (ou regressores)

$$\eta_i = \sum_{j=1}^p x_{ij}\beta_j = X\beta, \quad (\text{A.4})$$

Onde $X = (x_1, \dots, x_n)^T$ representa a matriz do modelo ; $\beta = (\beta_1, \dots, \beta_m)$ representa o vetor de parâmetros desconhecidos; $\eta = (\eta_1, \dots, \eta_m)^T$ representa o preditor linear.

- Uma função de ligação $g(\cdot)$ que é invertível, que transforma a esperança da variável resposta (μ_i) à um preditor linear,

$$\eta_i = g(\mu_i) = X\beta \quad (\text{A.5})$$

. Como a função de ligação $g(\cdot)$ é invertível, podemos escrever de seguinte forma

$$\mu_i = g^{-1}(\eta_i) = g^{-1}(X\beta). \quad (\text{A.6})$$

Considere a distribuição de Poisson com parâmetro $\lambda > 0$ usada para análise de dados na forma de contagem. A fmp é dada por

$$P(n, \lambda) = \frac{\lambda^n e^{-\lambda}}{n!} = \exp(n \log(\lambda) - \lambda - \log(n!)) \quad (\text{A.7})$$

. Assim, a distribuição de Poisson é membro da FE com $\theta = \log \lambda$, $b(\theta) = \exp(\theta)$; $C(n, \phi) = \log(n!)$, e $a(\phi) = 1$. A média e a variância de N são dadas por $E(N) = \dot{b}(\theta) = \lambda$ e, $\text{Var}(N) = a(\phi)V(\mu) = 1\lambda$, sendo $V(\mu) = \ddot{b}(\theta) = \lambda$ a função de variância.

Tabela A.1: Parâmetros da distribuição Poisson

θ	ϕ	$b(\theta)$	μ	$V(\mu)$	$C(n, \phi)$
$\log(\mu)$	1	$\mu = \exp(\theta)$	$\dot{b}(\theta) = \exp(\theta)$	$\ddot{b}(\theta) = \exp(\theta) = \mu$	$-\log(n!)$

A.2 Regressão Binomial

Assuma que N é uma variável aleatória com distribuição Binomial com fmp dada por

$$P(n, p) = \binom{m}{n} p^n (1-p)^{m-n}, \quad (\text{A.8})$$

com $n = 0, 1, 2, \dots, m$.

$$P(n; p) = \exp \left[n \log\left(\frac{p}{1-p}\right) + m \log(1-p) + \log \binom{m}{n} \right], \quad (\text{A.9})$$

. Desta forma, $P(n, p)$ é membro da FE com $\theta = \log\left(\frac{p}{1-p}\right)$; $b(\theta) = m \log(1 + \exp^\theta)$; $C(n, \phi) = \log \binom{m}{n}$, e $a(\phi) = 1$.

$$E(N) = \dot{b}(\theta) = mp \equiv \mu, \quad (\text{A.10})$$

$$V(\mu) = \frac{\partial \dot{b}(\theta)}{\partial \theta} = \frac{\partial}{\partial \theta} \left[n \frac{\exp(\theta)}{1 + \exp(\theta)} \right] = mp(1-p), \quad (\text{A.11})$$

$$\text{Var}(N) = a(\phi) \text{Var}(\mu) = mp(1 - p). \quad (\text{A.12})$$

sendo $\mu = mp$ a média da componente aleatória N .

Tabela A.2: Parâmetros da distribuição Binomial

θ	ϕ	$b(\theta)$	μ	$V(\mu)$	$C(y, \phi)$
$\log\left(\frac{\mu}{1-\mu}\right)$	1	$m \log(1 + e^\theta)$	$\frac{me^\theta}{1+e^\theta}$	$\mu(1 - \frac{\mu}{m})$	$\log\binom{m}{n}$

Apêndice B

Outras Distribuições de Contagens

B.1 Binomial Negativa

Binomial Negativa O modelo binomial negativo é uma distribuição de probabilidade discreta que descreve o número de tentativas necessárias para obter um certo número de sucessos em uma série de tentativas independentes. De acordo com Aryuyuen e Bodhisuwan (2013). A distribuição binomial negativa é frequentemente utilizada nos casos em que uma distribuição é sobredisperso. Se N representa uma variável aleatória distribuída segundo uma distribuição binomial negativa com parâmetro r e p , então a sua função de massa de probabilidade da distribuição binomial negativa é dada por

$$P(N = n) = \binom{r + n - 1}{n} p^r (1 - p)^n, \quad (\text{B.1})$$

em que $n = 0, 1, 2, \dots$ para $r > 0$ e $0 < p < 1$. Consequentemente obtemos

$$E[N] = \mu = \frac{r(1 - p)}{p}, \quad (\text{B.2})$$

$$E[N^2] = \frac{r(1 - p)}{(1 + r(1 - p))}, \quad (\text{B.3})$$

e

$$\text{Var}[N] = \frac{r(1-p)^2}{p}. \quad (\text{B.4})$$

O momento fatorial de N é

$$\mu_k(N) = E[n(N-1)\dots(N-k+1)] = \frac{\Gamma(r+k)(1-p)^k}{\Gamma(r)}, \quad (\text{B.5})$$

$k = 1, 2, \dots$, na qual $\Gamma(\cdot)$ é a função gama definida por

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx, \quad (\text{B.6})$$

em que n inteiros e $(\alpha, \beta > 0)$. O modelo binomial negativo, também conhecido como Poisson-Gama, é uma extensão do modelo Poisson que oferece uma maneira mais flexível de modelar dados com sobredispersão. Como sabemos a sobredispersão ocorre quando a variância dos dados é maior do que o esperado sob um modelo Poisson tradicional. Neste modelo, o parâmetro λ da distribuição de Poisson é tratado como uma variável aleatória que segue uma distribuição Gama com parâmetros $\Gamma(\alpha, \beta)$.

Seja $N|\lambda$ uma variável aleatória com distribuição de Poisson e λ tem distribuição *Gama* (α, β) , portanto

$$\begin{aligned} P(N = n) &= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^{\infty} \frac{e^{-\lambda}\lambda^n}{n!} \lambda^{\alpha-1} e^{-\lambda/\beta} d\lambda \\ &= \frac{1}{n!\Gamma(\alpha)\beta^\alpha} \int_0^{\infty} \lambda^{\alpha+n-1} e^{-\lambda(1+1/\beta)} d\lambda \\ &= \frac{1}{\Gamma(n+1)\Gamma(\alpha)\beta^\alpha} \frac{\Gamma(\alpha+n)}{\left(\frac{\beta}{\beta+1}\right)^{(\alpha+n)}} \\ &= \frac{\Gamma(\alpha+n-1)}{n} \left(1 + \frac{1}{\beta}\right)^\alpha \left(1 - \frac{1}{\beta+1}\right). \end{aligned} \quad (\text{B.7})$$

É notório que a distribuição marginal de N é Binomial Negativa com $r = \alpha$ e $p = \frac{1}{\beta+1}$. Se $r \rightarrow \infty$ e $p \rightarrow 1$ com média μ constante $P(N = n) \rightarrow \frac{e^{-\mu}\mu^n}{n!}$.

B.2 Binomial Negativa Inflacionado com Zeros

A distribuição binomial negativa inflacionada com zeros (DBNIZ), tem como parâmetros caraterísticos μ , ϕ e p , a sua expressão é dada da seguinte forma (Viviano, Muggero e Lovison, 2005)

$$f(n, p, \mu, \phi) = \begin{cases} p + (1 - p) \left(\frac{\phi}{\phi + \mu} \right)^\phi, & \text{se } n = 0, \\ (1 - p) \frac{\Gamma(n + \phi)}{\Gamma(\phi)n!} \left(\frac{\phi}{\phi + \mu} \right)^\phi \left(\frac{\mu}{\phi + \mu} \right)^n & \text{se } n > 0. \end{cases}$$

Da mesma maneira que o modelo Poisson inflacionado de zeros, a DBNIZ oferece uma estrutura para representar contagens nulas, referidas como estado zero, e para contagens não nulas, conhecidas como estado binomial negativo.

A média da DBNIZ e a sua variância podem ser encontradas de seguinte forma

$$\begin{aligned} E(N) &= \sum_{n=1}^{\infty} n f(n, p, \mu, \phi) \\ &= \sum_{n=1}^{\infty} n (1 - p) \frac{\Gamma(n + \phi)}{\Gamma(\phi)n!} \left(\frac{\phi}{\phi + \mu} \right)^\phi \left(\frac{\mu}{\phi + \mu} \right)^n \\ &= (1 - p) \frac{1}{\Gamma(\phi)} \left(\frac{\phi}{\phi + \mu} \right)^\phi \sum_{n=1}^{\infty} n \frac{\Gamma(n + \phi)}{n!} \frac{\mu^n}{(\phi + \mu)^n} \\ &= (1 - p) \frac{1}{\Gamma(\phi)} \left(\frac{\phi}{\phi + \mu} \right)^\phi \Gamma(1 + \phi) \mu \left(\frac{\phi}{\phi + \mu} \right)^{-\phi} \phi^{-1} \\ &= (1 - p) \mu. \end{aligned} \tag{B.8}$$

Agora vamos encontrar a variância, mas antes precisamos encontrar $E(N^2)$

$$\begin{aligned} E(N^2) &= \sum_{n=1}^{\infty} n^2 f(n, p, \mu, \phi) \\ &= (1 - p) \frac{1}{\Gamma(\phi)} \left(\frac{\phi}{\phi + \mu} \right)^\phi \sum_{n=1}^{\infty} n^2 \frac{\Gamma(n + \phi)}{n!} \frac{\mu^n}{(\phi + \mu)^n} \\ &= (1 - p) \frac{(\phi + (1 + \phi)\mu)\mu}{\phi} = (1 - p) \frac{(\phi + \mu + \mu\phi)}{\phi} \\ &= (1 - p) \mu \left(1 + \frac{\mu}{\phi} + \mu \right), \end{aligned} \tag{B.9}$$

então a $\text{Var}(N)$ é

$$\begin{aligned}
 \text{Var}(N) &= E(N^2) - [E(N)]^2 = (1-p)\mu \left(1 + \frac{\mu}{\phi} + \mu\right) - [(1-p)\mu]^2 \\
 &= (1-p)\mu \left(1 + \frac{\mu}{\phi} + \mu\right) - (\mu^2 - 2p\mu^2 + p^2\mu^2) \\
 &= \left(\mu + \frac{\mu^2}{\phi} + \mu^2 - p\mu - \frac{p\mu^2}{\phi} - p\mu^2\right) - (\mu^2 - 2p\mu^2 + p^2\mu^2) \\
 &= \mu \left(1 + \frac{\mu}{\phi} + p\mu - p - \frac{p\mu}{\phi} - p^2\mu\right) \\
 &= (1-p) \left(1 + \frac{\mu}{\phi} + p\mu\right) \mu. \tag{B.10}
 \end{aligned}$$

Os parâmetros μ , p e ϕ do modelo Binomial Negativo inflacionado de zeros são estimados com base na teoria de MLG. Para isso, o ajuste da modelo Binomial Negativo inflacionado de zeros, a ligação funcional depende de duas componentes

- A função de ligação da parte não inflacionada de zeros

$$\log(\mu) = X\beta, \tag{B.11}$$

- A função de ligação da parte inflacionada de zeros, considerando que o preditor linear é expresso pela função de ligação logit

$$\log\left(\frac{p}{1-p}\right) = G\gamma. \tag{B.12}$$

Onde X e G representam as matrizes de covariáveis para o modelo Binomial Negativo e para o modelo logístico, respectivamente, β e γ representam os vetores dos parâmetros dos modelos

Referências Bibliográficas

- Andrade, Bernardo B de et al. (2021). “On a weighted Poisson distribution and its associated regression model.” *Chilean Journal of Statistics (ChJS)* 12.2.
- Aryuyuen, Sirinapa e Bodhisuwan, Winai (2013). “The negative binomial-generalized exponential (NB-GE) distribution”. *Applied Mathematical Sciences* 7.22, pp. 1093–1105.
- Balakrishnan, N e Kozubowski, Tomasz J (2008). “A class of weighted Poisson processes”. *Statistics & probability letters* 78.15, pp. 2346–2352.
- Bhati, Deepesh, Sastry, D.V.S. e Qadri, P.Z. Maha (2015). “Uma Nova Distribuição Generalizada de Poisson-Lindley: Aplicações e Propriedades”. *Revista Austríaca de Estatística* 44.4, pp. 35–51.
- Boatwright, Peter et al. (2006). “Conjugate analysis of the Conway-Maxwell-Poisson distribution”.
- Castellares, Fredy, Lemonte, Artur J e Moreno-Arenas, Germán (2019). “On the two-parameter Bell–Touchard discrete distribution”. *Communications in Statistics-Theory and Methods* 49.19, pp. 4834–4852.
- Chrysaphinou, Ourania (1985). “On touchard polynomials”. *Discrete mathematics* 54.2, pp. 143–152.
- Consul, P. C. e Jain, G. C. (1973). “A Generalization of the Poisson Distribution”. *Technometrics* 15.4, pp. 791–799.

- Daley, DJ e Maindonald, JH (1989). “A unified view of models describing the avoidance of superparasitism”. *Mathematical Medicine and Biology: A Journal of the IMA* 6.3, pp. 161–178.
- Del Castillo, Joan e Pérez-Casany, Marta (1998). “Weighted Poisson distributions for overdispersion and underdispersion situations”. *Annals of the Institute of Statistical Mathematics* 50, pp. 567–585.
- ELAYDI, SABER e ZIEGLER, KLAUS (2004). “BERND AULBACH”, p. 349.
- Embrechts, Paul, Klüppelberg, Claudia e Mikosch, Thomas (2013). “Modelling extremal events: for insurance and finance”. 33.
- Gross, Donald et al. (2011). “Fundamentals of queueing theory”. 627.
- Hall, Daniel B (2000). “Zero-inflated Poisson and binomial regression with random effects: a case study”. *Biometrics* 56.4, pp. 1030–1039.
- Ho, Linda L et al. (2021). “Monitoring count data with Shewhart control charts based on the Touchard model”. *Quality and Reliability Engineering International* 37.5, pp. 1875–1893.
- Johnson, Norman L, Kemp, Adrienne W e Kotz, Samuel (2005). *Univariate discrete distributions*. Vol. 444. John Wiley & Sons.
- Kokonendji, Célestin C, Mizere, Dominique e Balakrishnan, Narayanaswamy (2008). “Connections of the Poisson weight function to overdispersion and underdispersion”. *Journal of Statistical Planning and Inference* 138.5, pp. 1287–1296.
- Lambert, Diane (1992). “Zero-inflated Poisson regression, with an application to defects in manufacturing”. *Technometrics* 34.1, pp. 1–14.
- Lerner, Bores, Lone, Amjad e Rao, Murali (1997). “On generalized Poisson distributions”. *PROBABILITY AND MATHEMATICAL STATISTICS-WROCLAW UNIVERSITY* 17, pp. 377–385.
- Magalhães, Marcos Nascimento (2006). *Probabilidade e variáveis aleatórias*. Edusp.
- Matsushita, Raul e Oliveira, Sandro (2016). “A Distribuição Touchard e suas Aplicações”, p. 73.

- Matsushita, Raul et al. (2019). “The Touchard distribution”. *Communications in Statistics-Theory and Methods* 48.8, pp. 2049–2059.
- Meyer, Paul L (1983). *Probabilidade: aplicações à estatística*. Livros Técnicos e Científicos Rio de Janeiro.
- Mijburgh, Philip Albert (2020). “On weighted Poisson distributions and processes, with associated inference and applications”. Tese de dout. University of Pretoria (South Africa).
- Mizère, D, Kokonendji, CC e Dossou-Gbété, Simplicie (2006). “Quelques tests de la loi de Poisson contre des alternatives générales basées sur l’indice de dispersion de Fisher”. *Revue de statistique appliquée* 54.4, pp. 61–84.
- Mizere, Dominique et al. (2013). “Décomposition d’une loi de poisson pondérée en une combinaison convexe de lois duales”. *Afrika Statistika* 8.1, pp. 583–594.
- Nelder, John Ashworth e Wedderburn, Robert WM (1972). “Generalized linear models”. *Journal of the Royal Statistical Society Series A: Statistics in Society* 135.3, pp. 370–384.
- Nkounkou, CL et al. (2017). “Approvisionnement en eau dans la ville de Brazzaville, Congo”. *European Scientific Journal* 13.21, pp. 474–498.
- Patil, Ganapati P e Rao, Calyampudi R (1978). “Weighted distributions and size-biased sampling with applications to wildlife populations and human families”. *Biometrics*, pp. 179–189.
- Ross, Sheldon M (2014). “Introduction to probability models”.
- Shanker, Rama, Sharma, Shambhu e Shanker, Ravi (2013). “A two-parameter Lindley distribution for modeling waiting and survival times data”.
- Shmueli, Galit et al. (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, pp. 127–142.
- Snyder, Conway W, Neugebauer, Marcia e Rao, UR (1963). “The solar wind velocity and its correlation with cosmic-ray variations and with solar and geomagnetic activity”. *Journal of Geophysical Research* 68.24, pp. 6361–6370.

§B.2. REFERÊNCIAS BIBLIOGRÁFICAS

Touchard, Jacques (1933). *Propriétés arithmétiques de certains nombres récurrents*. Secrétariat de la société scientifique.

Viviano, Lorena, Muggero, Vito MR e Lovison, Gianfranco (2005). “Using Zero-inflated Models to Analyze Environmental Data Sets with Many Zeroes”.