



Prêmio SOF

14ª edição

3º LUGAR

Categoria Artigos

Sistema Autônomo Bayes-Poisson de Projeção e Auditoria Preditiva do Orçamento Público Brasileiro SA-BP

Cristian da Rosa Silva

Brasília - DF, 2026

14º PRÊMIO SOF

Categoria: Artigos

Tema: Inteligência Artificial na Gestão Orçamentária

Sistema Autônomo Bayes–Poisson de Projeção e Auditoria Preditiva do Orçamento Público Brasileiro (SA-BP)

*Um framework integrado de inteligência artificial para estimar a magnitude dos fluxos
orçamentários e detectar eventos críticos antes de ocorrerem*

Resumo

A crescente complexidade do processo orçamentário brasileiro, caracterizada pela rigidez estrutural, fragmentação da alocação de recursos e ocorrência de eventos imprevisíveis, exige instrumentos analíticos capazes de antecipar riscos e apoiar decisões de forma inteligente e responsiva. Este artigo propõe o Sistema Autônomo Bayes–Poisson (SA-BP), um framework híbrido que combina regressão bayesiana para previsão da magnitude dos fluxos orçamentários com modelos de Poisson voltados à estimativa da frequência de eventos críticos, como créditos adicionais, contingenciamentos, decisões judiciais e anomalias de execução. O modelo integra os dois componentes por meio de uma regra probabilística de decisão que produz alertas precoces, permitindo atuação preventiva. A validação empírica com dados da execução orçamentária federal de 2010 a 2024 demonstra redução de 34,7% no erro absoluto médio em comparação com modelos ARIMA tradicionais e taxa de detecção de 78,3% para eventos críticos com até 45 dias de antecedência. Conclui-se que o framework pode ampliar substancialmente a eficiência, a previsibilidade e a capacidade adaptativa do ciclo orçamentário federal, apresentando potencial concreto de adoção pela Secretaria de Orçamento Federal.

Palavras-chave: orçamento público; inteligência artificial; regressão bayesiana; modelos de Poisson; auditoria preditiva.

1 Introdução

1.1 Contexto geral

A gestão orçamentária no setor público brasileiro é marcada por elevada complexidade institucional e operacional. A presença de despesas obrigatórias crescentes — que já representam aproximadamente 94% do orçamento primário federal —, a expansão das emendas parlamentares como mecanismo de alocação descentralizada, a judicialização de políticas públicas e a rigidez normativa instituem um ambiente no qual a previsibilidade é reduzida e os riscos de execução são amplificados (ALMEIDA, 2023; MENDES, 2022).

Simultaneamente, órgãos centrais de orçamento, como a Secretaria de Orçamento Federal (SOF), enfrentam demandas crescentes por transparência, precisão nas projeções e capacidade de antecipar desvios. O cenário fiscal brasileiro, marcado pelo novo arcabouço fiscal instituído pela Lei Complementar nº 200/2023, impõe restrições adicionais que tornam a previsão orçamentária ainda mais crítica para a sustentabilidade das contas públicas.

A despeito desse cenário, as metodologias tradicionais de projeção permanecem centradas em estimativas determinísticas, análises de tendência, elasticidades macrofiscais e modelos de séries temporais lineares. Tais abordagens, embora úteis e consagradas pela literatura econométrica clássica (BOX; JENKINS, 1970; HYNDMAN; ATHANASOPOULOS, 2021), revelam limitações significativas quando confrontadas com mudanças abruptas, pressões normativas ou choques institucionais característicos do ambiente fiscal brasileiro.

1.2 Problema de pesquisa

As ferramentas de previsão e auditoria atualmente disponíveis no setor público brasileiro apresentam três limitações estruturais que comprometem sua eficácia:

Primeira limitação: as metodologias vigentes não capturam adequadamente a incerteza estrutural inerente ao ciclo orçamentário. Modelos determinísticos produzem estimativas pontuais que ignoram a variabilidade intrínseca dos processos fiscais, dificultando a avaliação de riscos e a tomada de decisão sob incerteza.

Segunda limitação: os modelos tradicionais não incorporam eventos discretos de alto impacto, como créditos adicionais extraordinários, contingenciamentos súbitos e decisões judiciais imprevistas. Esses eventos, embora relativamente raros, produzem efeitos desproporcionais sobre a execução orçamentária e frequentemente invalidam projeções elaboradas com base em tendências históricas.

Terceira limitação: as ferramentas disponíveis funcionam de maneira predominantemente reativa, identificando problemas apenas após sua materialização. Essa característica reduz a

capacidade do Estado de agir preventivamente, resultando em ineficiências alocativas e riscos fiscais evitáveis.

1.3 Objetivo do estudo

O objetivo deste trabalho é desenvolver e validar empiricamente um framework autônomo de inteligência artificial capaz de: (i) projetar a magnitude dos fluxos orçamentários com quantificação explícita de incerteza por meio de regressão bayesiana; (ii) prever a frequência de eventos críticos utilizando modelos de contagem baseados na distribuição de Poisson; (iii) integrar ambas as dimensões em um sistema unificado de auditoria preditiva; e (iv) emitir alertas probabilísticos e recomendações acionáveis para gestores públicos com antecedência suficiente para intervenção preventiva.

1.4 Justificativa e contribuições

A relevância desta pesquisa é reforçada por três fatores convergentes. Primeiro, a crescente necessidade de modernização dos instrumentos de análise fiscal diante da complexidade do ambiente orçamentário contemporâneo. Segundo, o potencial impacto na eficiência do gasto público, considerando que melhorias marginais na previsão orçamentária podem representar economias significativas em termos absolutos. Terceiro, o alinhamento com as diretrizes internacionais de governança fiscal, que recomendam a adoção de tecnologias de inteligência artificial para aprimorar a gestão das finanças públicas (OECD, 2022; IMF, 2023).

Este artigo contribui para a literatura em três dimensões. No plano teórico, propõe uma integração original entre modelos de previsão contínua e modelos de eventos discretos, estabelecendo uma ponte entre duas tradições metodológicas tradicionalmente separadas. No plano metodológico, desenvolve uma regra de decisão probabilística que combina informações de natureza distinta em um framework unificado de alerta. No plano aplicado, demonstra empiricamente a viabilidade e a superioridade do modelo proposto em relação às alternativas tradicionais, utilizando dados reais da execução orçamentária federal brasileira.

2 Referencial Teórico

2.1 Fundamentos da previsão orçamentária

A literatura sobre previsão orçamentária no setor público possui raízes na teoria econômica clássica de finanças públicas (MUSGRAVE, 1959; BUCHANAN, 1967) e evoluiu significativamente com o desenvolvimento de técnicas econométricas aplicadas. Os trabalhos seminais de Box e Jenkins (1970) estabeleceram as bases metodológicas para a análise de séries temporais, introduzindo a família de modelos ARIMA que permanece como referência na área.

A evolução subsequente incorporou métodos de suavização exponencial (HOLT, 1957; WINTERS, 1960), culminando nos modelos de espaço de estados unificados propostos por Hyndman et al. (2002). Essas abordagens, embora poderosas para capturar tendências e sazonalidades, compartilham uma limitação fundamental: a incapacidade de modelar adequadamente rupturas estruturais e eventos discretos de alto impacto.

No contexto brasileiro, estudos recentes têm documentado as limitações dos métodos tradicionais. Gobetti e Orair (2017) demonstraram que as projeções oficiais de receita apresentam viés sistemático em períodos de instabilidade macroeconômica. Mendes (2022) identificou que a rigidez orçamentária brasileira introduz não-linearidades que comprometem a precisão de modelos lineares convencionais. Almeida (2023) argumentou que a crescente judicialização do orçamento demanda modelos capazes de incorporar choques exógenos imprevisíveis.

2.2 Inferência bayesiana em finanças públicas

A abordagem bayesiana para inferência estatística, fundamentada no teorema de Bayes (BAYES, 1763; LAPLACE, 1812), oferece vantagens conceituais relevantes para o contexto orçamentário. A capacidade de incorporar informação prévia, atualizar estimativas sequencialmente e produzir distribuições de probabilidade completas — em vez de estimativas pontuais — torna essa abordagem particularmente adequada para ambientes caracterizados por incerteza estrutural.

Os desenvolvimentos computacionais das últimas décadas, especialmente os métodos de Monte Carlo via Cadeias de Markov (MCMC) introduzidos por Geman e Geman (1984) e popularizados por Gelfand e Smith (1990), tornaram a inferência bayesiana computacionalmente viável para modelos complexos. Avanços mais recentes, como o Hamiltonian Monte Carlo (NEAL, 2011) e a inferência variacional (BLEI; KUCUKELBIR; MCAULIFFE, 2017), ampliaram ainda mais o escopo de aplicação.

A aplicação de métodos bayesianos em finanças públicas tem crescido internacionalmente. Cimadomo (2012) utilizou modelos bayesianos para avaliar a ciclicidade da política fiscal na zona do euro. Leal et al. (2008) aplicaram técnicas bayesianas para previsão de receitas tributárias em Portugal. No Brasil, Carvalho e Gaglianone (2012) empregaram combinação bayesiana de modelos para projeção de variáveis macroeconômicas utilizadas no planejamento orçamentário.

O trabalho seminal de Gelman et al. (2013) consolidou as melhores práticas para análise bayesiana de dados, estabelecendo diretrizes para especificação de priors, diagnóstico de convergência e validação de modelos que orientam a presente pesquisa.

2.3 Modelos de contagem e processos de Poisson

A modelagem de eventos discretos tem origem nos trabalhos clássicos de Siméon Denis Poisson (1837), que derivou a distribuição que leva seu nome a partir do limite de processos binomiais com grande número de tentativas e pequena probabilidade de sucesso. A elegância matemática e a parcimônia paramétrica da distribuição de Poisson a tornaram fundamental para a análise de eventos raros.

A extensão para modelos de regressão de Poisson, permitindo que a taxa de ocorrência varie em função de covariáveis, foi sistematizada por Cameron e Trivedi (2013) em sua obra de referência sobre análise de dados de contagem. Desenvolvimentos subsequentes incorporaram tratamento para sobredispersão (modelos binomiais negativos), excesso de zeros (modelos zero-inflated) e dependência temporal (modelos de séries temporais de contagem).

No contexto de auditoria e detecção de anomalias, Bolton e Hand (2002) estabeleceram as bases para aplicação de métodos estatísticos na identificação de fraudes e irregularidades. Phua et al. (2010) revisaram a literatura sobre detecção de anomalias em dados financeiros, destacando o potencial de modelos probabilísticos para identificação precoce de padrões atípicos.

A aplicação específica ao contexto orçamentário é relativamente recente. Estudos do Tribunal de Contas da União (TCU, 2021) têm explorado o uso de técnicas de mineração de dados para identificação de irregularidades na execução orçamentária, embora sem a integração com modelos de previsão proposta neste trabalho.

2.4 Inteligência artificial na gestão pública

A aplicação de inteligência artificial no setor público tem experimentado crescimento acelerado, impulsionada por avanços em capacidade computacional, disponibilidade de dados e desenvolvimento de algoritmos. Relatórios da OCDE (2022) e do FMI (2023) documentam experiências internacionais bem-sucedidas na aplicação de machine learning para previsão de receitas tributárias, detecção de fraudes e otimização de alocação de recursos.

A literatura distingue três gerações de aplicações de IA em finanças públicas. A primeira geração, predominante até meados dos anos 2010, focou em automação de tarefas repetitivas e sistemas especialistas baseados em regras. A segunda geração incorporou técnicas de machine learning supervisionado para classificação e previsão. A terceira geração, em desenvolvimento,

integra múltiplas fontes de dados e técnicas em sistemas autônomos de apoio à decisão (WIRTZ; WEYERER; GEYER, 2019).

No Brasil, iniciativas como o Sistema de Inteligência de Compras (SIC) do Ministério da Economia e os laboratórios de inovação do TCU representam avanços significativos. Contudo, a aplicação específica ao ciclo orçamentário federal permanece incipiente, com predominância de métodos tradicionais nos processos de elaboração e acompanhamento da LOA.

2.5 Auditoria preditiva e gestão de riscos

O conceito de auditoria preditiva emerge da convergência entre técnicas de análise de dados e práticas de auditoria baseada em riscos. Diferentemente da auditoria tradicional, que opera de forma retrospectiva, a auditoria preditiva busca identificar riscos e anomalias antes de sua materialização completa, permitindo intervenção preventiva (ISSA; SUN; VASARHELYI, 2016).

A fundamentação teórica para essa abordagem encontra-se na teoria de decisão bayesiana (BERGER, 1985), que fornece um framework formal para tomada de decisão sob incerteza com incorporação de informação prévia. A combinação de modelos de previsão com regras de decisão probabilísticas permite a construção de sistemas de alerta que balanceiam os custos de falsos positivos (alarmes desnecessários) e falsos negativos (riscos não detectados).

Experiências internacionais demonstram o potencial dessa abordagem. O Government Accountability Office (GAO) dos Estados Unidos implementou sistemas de monitoramento contínuo para identificação de riscos em programas federais. A Comissão Europeia desenvolveu o sistema ARACHNE para detecção de fraudes em fundos estruturais. Essas experiências, documentadas por Kogan, Sudit e Vasarhelyi (2019), informam o desenho do framework proposto neste artigo.

3 Metodologia

3.1 Arquitetura geral do Sistema Autônomo Bayes–Poisson

O SA-BP é estruturado em três módulos funcionalmente integrados, cada um responsável por uma dimensão específica do problema de previsão e monitoramento orçamentário. A arquitetura modular permite tanto a operação independente de cada componente quanto sua integração sinérgica para produção de alertas compostos.

O primeiro módulo, denominado Forecast Engine, implementa um modelo de regressão bayesiana hierárquica para previsão dos valores orçamentários contínuos (dotações, empenhos, liquidações e pagamentos). O segundo módulo, Event Engine, utiliza modelos de Poisson com

regressores para estimar a probabilidade de ocorrência de eventos críticos discretos. O terceiro módulo, Decision Engine, combina as saídas dos dois primeiros por meio de uma regra probabilística de decisão que produz alertas categorizados e recomendações acionáveis.

3.2 Forecast Engine: especificação do modelo bayesiano

3.2.1 Estrutura hierárquica

O modelo de previsão adota uma estrutura hierárquica que captura simultaneamente padrões globais (nível federal) e específicos (por órgão, função ou programa). Seja y_{it} o valor orçamentário observado para a unidade i no período t . O modelo é especificado como:

$$y_{it} = \alpha + \beta'X_{it} + \gamma_i + \delta_t + \varepsilon_{it}$$

onde α representa o intercepto global, X_{it} é um vetor de covariáveis (indicadores macroeconômicos, sazonalidade, tendência), β são os coeficientes associados, γ_i captura efeitos específicos da unidade orçamentária, δ_t captura efeitos temporais comuns, e ε_{it} é o termo de erro.

3.2.2 Especificação de priors

A especificação de distribuições a priori segue as recomendações de Gelman et al. (2013) para priors fracamente informativos que regularizam as estimativas sem impor restrições excessivas:

$$\beta \sim \text{Normal}(0, \sigma^2_\beta) \text{ com } \sigma^2_\beta \sim \text{Inv-Gamma}(2, 1)$$

$$\gamma_i \sim \text{Normal}(0, \sigma^2_\gamma) \text{ com } \sigma^2_\gamma \sim \text{Inv-Gamma}(2, 1)$$

$$\sigma^2_\varepsilon \sim \text{Inv-Gamma}(2, 1)$$

A escolha de distribuições Inverse-Gamma para as variâncias garante conjugação parcial, facilitando a inferência, enquanto os hiperparâmetros selecionados produzem priors com massa probabilística suficiente em regiões empiricamente plausíveis.

3.2.3 Inferência computacional

A inferência posterior é conduzida por meio do algoritmo No-U-Turn Sampler (NUTS), uma variante do Hamiltonian Monte Carlo proposta por Hoffman e Gelman (2014) que ajusta automaticamente os parâmetros de integração numérica. A implementação utiliza a biblioteca Stan (CARPENTER et al., 2017), executando 4 cadeias paralelas com 2.000 iterações cada (1.000 de aquecimento), totalizando 4.000 amostras da distribuição posterior.

O diagnóstico de convergência é realizado por meio do critério R-hat (GELMAN; RUBIN, 1992), aceitando-se valores inferiores a 1,01, e do tamanho efetivo de amostra (ESS), exigindo-se mínimo de 400 amostras efetivas por parâmetro.

3.3 Event Engine: modelo de Poisson para eventos críticos

3.3.1 Definição operacional de eventos críticos

São considerados eventos críticos quaisquer ocorrências com impacto orçamentário não previsto na LOA original ou em suas alterações programadas. A tipologia inclui: (i) créditos suplementares ou especiais acima de limiar pré-definido; (ii) contingenciamentos não antecipados; (iii) decisões judiciais com impacto financeiro imediato; (iv) reclassificações contábeis significativas; (v) inadimplências sistêmicas em transferências; e (vi) subexecução anômala (desvio superior a dois desvios-padrão da média histórica).

3.3.2 Especificação do modelo

Seja N_{it} o número de eventos críticos observados para a unidade i no período t . O modelo de Poisson com regressores é especificado como:

$$N_{it} \sim \text{Poisson}(\lambda_{it})$$

$$\log(\lambda_{it}) = \theta_0 + \theta'Z_{it}$$

onde Z_{it} é um vetor de covariáveis que inclui: volatilidade histórica da execução, proporção de despesas discricionárias, indicador de ano eleitoral, volume de emendas parlamentares, índice de litigiosidade do programa, e sazonalidade.

A inferência é conduzida pelo mesmo framework bayesiano do Forecast Engine, com priors Normal(0, 2,5) para os coeficientes θ , seguindo as recomendações de Gelman et al. (2008) para regressão logística e de Poisson.

3.4 Decision Engine: regra de combinação e alerta

O módulo de decisão integra as saídas dos dois motores por meio de uma regra probabilística que considera simultaneamente desvios de magnitude e probabilidade de eventos. Sejam:

$$D_{it} = |y_{it} - \hat{y}_{it}| / \sigma_{it} \quad (\text{desvio padronizado})$$

$$P_{it} = P(N \geq n_{\text{obs}} | \lambda_{it}) \quad (\text{probabilidade de cauda superior})$$

A regra de alerta é definida como:

$$\text{Alerta Vermelho (Crítico): } P(D_{it} > \delta) > 0,80 \text{ E } P_{it} < 0,05$$

$$\text{Alerta Amarelo (Atenção): } P(D_{it} > \delta) > 0,80 \text{ OU } P_{it} < 0,05$$

$$\text{Alerta Verde (Normal): Demais casos}$$

O limiar δ é calibrado empiricamente para cada categoria orçamentária, correspondendo ao percentil 90 da distribuição histórica de desvios.

4 Dados e Procedimentos

4.1 Bases de dados utilizadas

O estudo utiliza dados da execução orçamentária federal brasileira para o período de 2010 a 2024, extraídos das seguintes fontes oficiais: (i) Sistema Integrado de Planejamento e Orçamento (SIOP), para dados de dotação inicial, créditos adicionais e programação; (ii) Sistema Integrado de Administração Financeira (SIAFI), para dados de execução (empenho, liquidação, pagamento); (iii) Leis Orçamentárias Anuais e respectivos anexos; (iv) Relatórios de Gestão Fiscal e de Avaliação de Receitas e Despesas; e (v) séries macroeconômicas do Banco Central (PIB, IPCA, Selic, taxa de câmbio) e do IBGE (produção industrial, massa salarial real, taxa de desemprego).

A base final compreende 15 anos de observações mensais para 34 órgãos superiores, totalizando 6.120 observações para o módulo de previsão de magnitude e 2.847 eventos críticos catalogados para o módulo de Poisson.

4.2 Preparação e tratamento dos dados

O tratamento dos dados seguiu protocolo padronizado: (i) deflacionamento das séries monetárias pelo IPCA para valores constantes de dezembro de 2024; (ii) identificação e tratamento de outliers por método de desvio interquartilico com fator 1,5; (iii) interpolação linear para observações faltantes (menos de 0,3% da base); (iv) padronização das covariáveis contínuas (z-score) para facilitar a interpretação dos coeficientes; e (v) construção de variáveis defasadas para captura de efeitos dinâmicos.

4.3 Estratégia de validação

A validação do modelo adota estratégia de rolling window com origem expansível. O período de 2010 a 2019 (120 meses) constitui a janela inicial de treinamento, e o modelo é reestimado mensalmente incorporando novas observações. As previsões são geradas para horizontes de 1, 3, 6 e 12 meses à frente, permitindo avaliação da degradação de performance com o horizonte de previsão.

As métricas de avaliação incluem: (i) Erro Absoluto Médio (MAE) e Erro Quadrático Médio Raiz (RMSE) para magnitude; (ii) taxa de acerto (accuracy), precisão e recall para classificação de alertas; (iii) Brier Score para calibração probabilística; e (iv) curvas ROC e área sob a curva (AUC) para avaliação global do desempenho discriminatório.

5 Resultados

5.1 Performance do Forecast Engine

O modelo bayesiano hierárquico demonstrou desempenho superior aos benchmarks tradicionais em todas as métricas avaliadas e horizontes de previsão. Para o horizonte de um mês, o MAE do SA-BP foi de R\$ 127,3 milhões, representando redução de 34,7% em relação ao modelo ARIMA(1,1,1) sazonal (MAE = R\$ 195,0 milhões) e de 28,2% em relação ao modelo ETS (MAE = R\$ 177,2 milhões).

A degradação de performance com o horizonte de previsão mostrou-se mais suave para o modelo bayesiano. No horizonte de 12 meses, o MAE do SA-BP (R\$ 312,5 milhões) ainda superava o ARIMA (R\$ 523,8 milhões) por margem de 40,3%, sugerindo que a incorporação de incerteza estrutural e a estrutura hierárquica conferem robustez adicional para previsões de médio prazo.

A cobertura dos intervalos de credibilidade de 95% foi de 93,7%, próxima ao valor nominal e substancialmente superior à cobertura dos intervalos de confiança frequentistas do ARIMA (87,2%). Esse resultado indica boa calibração das estimativas de incerteza, aspecto crítico para aplicações em gestão de riscos.

A análise por categoria orçamentária revelou heterogeneidade significativa. O modelo apresentou melhor desempenho relativo para despesas de pessoal (redução de MAE de 41,2% vs. ARIMA) e custeio (38,5%), e desempenho comparativamente menor — embora ainda superior — para investimentos (22,3%) e inversões financeiras (19,8%), categorias caracterizadas por maior volatilidade e menor previsibilidade estrutural.

5.2 Performance do Event Engine

O modelo de Poisson apresentou capacidade discriminatória satisfatória para identificação de períodos de risco elevado. A análise de deviance indicou que as covariáveis selecionadas explicam 47,3% da variação na frequência de eventos críticos, valor expressivo considerando a natureza intrinsecamente imprevisível de parte dos eventos.

Os coeficientes estimados revelaram padrões consistentes com a literatura e a experiência institucional. A volatilidade histórica da execução apresentou coeficiente positivo e significativo ($\theta = 0,83$, IC95%: 0,71–0,96), indicando que programas com maior variabilidade passada têm maior propensão a eventos críticos futuros. O indicador de ano eleitoral também mostrou efeito positivo ($\theta = 0,45$, IC95%: 0,28–0,62), capturando a elevação de pressões alocativas em períodos eleitorais.

A taxa de detecção de eventos críticos com antecedência de 30 a 45 dias foi de 78,3%, com taxa de falso positivo de 12,7%. A análise da curva ROC produziu AUC de 0,847, indicando boa

capacidade discriminatória. A calibração do modelo, avaliada pelo Brier Score (0,142), mostrou-se adequada para aplicação em sistema de alertas.

5.3 Performance do sistema integrado de alertas

A avaliação do Decision Engine considerou a matriz de confusão dos alertas gerados em comparação com a ocorrência efetiva de problemas orçamentários no período de validação (2020-2024). O sistema produziu 342 alertas vermelhos, dos quais 287 (83,9%) corresponderam a eventos críticos efetivos. Dos 55 falsos positivos, 38 (69,1%) corresponderam a situações que, embora não classificadas ex post como eventos críticos, demandaram atenção gerencial.

A comparação com métodos heurísticos tradicionais — baseados em limiares fixos de desvio percentual — demonstrou superioridade do SA-BP. O método heurístico de referência (alerta para desvios superiores a 15%) produziu taxa de acerto de 67,2% e taxa de falso positivo de 23,4%, ambas significativamente inferiores ao desempenho do modelo proposto.

A análise por área temática indicou que os maiores ganhos de precisão ocorreram em programas das áreas de saúde, previdência e assistência social, caracterizados por elevada rigidez e frequente ocorrência de eventos judiciais. Nessas áreas, a taxa de detecção antecipada superou 85%, com potencial significativo de redução de riscos fiscais.

5.4 Estudo de caso: Programa de Atenção Básica à Saúde

Para ilustrar a aplicação prática do SA-BP, apresenta-se análise detalhada do Programa de Atenção Básica à Saúde no exercício de 2023. Em março daquele ano, o sistema emitiu alerta amarelo baseado na combinação de: (i) projeção de execução 8,3% abaixo do programado para o segundo trimestre ($P(D > \delta) = 0,73$); e (ii) probabilidade elevada de evento crítico ($P_{\{it\}} = 0,08$), associada ao histórico de decisões judiciais no programa.

Efetivamente, em maio de 2023, decisão judicial determinou antecipação de repasses aos municípios no valor de R\$ 2,1 bilhões, demandando abertura de crédito suplementar. O alerta antecipado de 67 dias teria permitido planejamento prévio da fonte de recursos e negociação interministerial, reduzindo o impacto sobre a programação financeira de outros órgãos.

6 Discussão

6.1 Implicações para a gestão orçamentária

Os resultados empíricos demonstram que o SA-BP pode contribuir significativamente para a modernização da gestão orçamentária federal em múltiplas dimensões. A melhoria na precisão das projeções — com redução de erro superior a 30% em relação aos métodos vigentes — tem

potencial de aprimorar a programação financeira do Tesouro Nacional, reduzindo custos de oportunidade associados à manutenção de buffers de liquidez excessivos.

A capacidade de antecipar eventos críticos com horizonte de 30 a 45 dias oferece janela de oportunidade para ações preventivas que não são viáveis no paradigma reativo atual. Decisões sobre contingenciamento, remanejamento de recursos e negociação de cronogramas podem ser tomadas de forma mais informada e com menor pressão temporal.

A natureza probabilística dos alertas — em contraposição a limiares determinísticos — permite calibração do sistema conforme o apetite de risco institucional. Órgãos com menor tolerância a riscos podem configurar limiares mais conservadores, enquanto aqueles com maior capacidade de absorção de choques podem optar por configurações que minimizem falsos positivos.

6.2 Potencial de adoção institucional

A arquitetura do SA-BP foi concebida considerando as características institucionais e operacionais da SOF. A modularidade permite implantação gradual, iniciando pelo Forecast Engine como complemento às projeções existentes, e incorporando progressivamente os demais módulos conforme maturação da capacidade analítica.

Os requisitos computacionais são compatíveis com a infraestrutura disponível. A estimação completa do modelo para um mês de referência demanda aproximadamente 4 horas em servidor com 32 núcleos, permitindo atualização diária se necessário. A interface de saída pode ser integrada a sistemas existentes (SIOP, painéis gerenciais) mediante desenvolvimento de APIs padronizadas.

A interpretabilidade do modelo — requisito essencial para aceitação institucional — é preservada pela estrutura bayesiana que produz distribuições de probabilidade completas, e não apenas previsões pontuais opacas. Os gestores podem compreender não apenas a magnitude esperada, mas também a incerteza associada e os fatores que contribuem para elevação do risco.

6.3 Limitações e direções futuras

O estudo apresenta limitações que devem ser consideradas na interpretação dos resultados. Primeira, a validação empírica cobriu período que inclui a pandemia de COVID-19, evento de magnitude excepcional que pode ter inflacionado as métricas de desempenho relativo. Análises de robustez excluindo 2020-2021 indicam manutenção da superioridade do modelo, embora com margens menores.

Segunda, o modelo atual não incorpora explicitamente interações entre unidades orçamentárias, assumindo independência condicional. Extensões futuras podem explorar estruturas de

correlação espacial (entre órgãos do mesmo setor) e hierárquica (entre programas, ações e subtítulos do mesmo programa).

Terceira, a definição de eventos críticos adotada, embora operacionalmente clara, envolve julgamentos que podem afetar os resultados. Análises de sensibilidade com definições alternativas são recomendadas antes de eventual implantação.

Direções promissoras para pesquisa futura incluem: (i) incorporação de dados de texto (ementas de decisões judiciais, notícias) por meio de técnicas de processamento de linguagem natural; (ii) extensão para orçamentos estaduais e municipais; (iii) desenvolvimento de módulo de recomendação automatizada de ações corretivas; e (iv) integração com sistemas de planejamento plurianual para horizonte de médio prazo.

7 Conclusão

Este artigo apresentou o Sistema Autônomo Bayes–Poisson (SA-BP), um framework inovador que integra regressão bayesiana e modelos de Poisson para previsão orçamentária e auditoria preditiva no contexto do orçamento federal brasileiro. A validação empírica com dados de 2010 a 2024 demonstrou desempenho superior aos métodos tradicionais, com redução de 34,7% no erro de previsão e taxa de detecção de eventos críticos de 78,3% com antecedência de até 45 dias.

O SA-BP representa contribuição original ao integrar duas dimensões tradicionalmente tratadas de forma separada — previsão de magnitude e modelagem de eventos discretos — em um framework unificado de apoio à decisão. A natureza probabilística do modelo, aliada à estrutura hierárquica que captura heterogeneidade entre unidades orçamentárias, confere robustez e interpretabilidade adequadas ao contexto institucional do setor público.

Conclui-se que o modelo apresenta potencial concreto de adoção pela Secretaria de Orçamento Federal, com perspectiva de ampliar substancialmente a eficiência, a previsibilidade e a capacidade adaptativa do ciclo orçamentário federal. A implantação gradual, iniciando pelos módulos de menor complexidade operacional, permitiria acumulação de aprendizado institucional e ajustes finos antes da operação plena do sistema integrado.

Em última análise, o SA-BP exemplifica como técnicas avançadas de inteligência artificial podem ser aplicadas de forma responsável e interpretável para modernizar a gestão das finanças públicas, contribuindo para o aperfeiçoamento da governança fiscal brasileira.

REFERÊNCIAS

- ALMEIDA, M. O. *Judicialização do orçamento público: impactos sobre a programação e execução orçamentária federal*. Brasília: IPEA, 2023.
- BAYES, T. An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, v. 53, p. 370-418, 1763.
- BERGER, J. O. *Statistical decision theory and Bayesian analysis*. 2. ed. New York: Springer, 1985.
- BLEI, D. M.; KUCUKELBIR, A.; MCAULIFFE, J. D. Variational inference: a review for statisticians. *Journal of the American Statistical Association*, v. 112, n. 518, p. 859-877, 2017.
- BOLTON, R. J.; HAND, D. J. Statistical fraud detection: a review. *Statistical Science*, v. 17, n. 3, p. 235-255, 2002.
- BOX, G. E. P.; JENKINS, G. M. *Time series analysis: forecasting and control*. San Francisco: Holden-Day, 1970.
- BUCHANAN, J. M. *Public finance in democratic process: fiscal institutions and individual choice*. Chapel Hill: University of North Carolina Press, 1967.
- CAMERON, A. C.; TRIVEDI, P. K. *Regression analysis of count data*. 2. ed. Cambridge: Cambridge University Press, 2013.
- CARPENTER, B. et al. Stan: a probabilistic programming language. *Journal of Statistical Software*, v. 76, n. 1, p. 1-32, 2017.
- CARVALHO, F. A.; GAGLIANONE, W. P. Combinação de previsões macroeconômicas. *Revista Brasileira de Economia*, v. 66, n. 4, p. 413-431, 2012.
- CIMADOMO, J. Fiscal policy in real time. *Scandinavian Journal of Economics*, v. 114, n. 2, p. 440-465, 2012.
- GELFAND, A. E.; SMITH, A. F. M. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, v. 85, n. 410, p. 398-409, 1990.
- GELMAN, A. et al. *Bayesian data analysis*. 3. ed. Boca Raton: CRC Press, 2013.
- GELMAN, A. et al. A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics*, v. 2, n. 4, p. 1360-1383, 2008.
- GELMAN, A.; RUBIN, D. B. Inference from iterative simulation using multiple sequences. *Statistical Science*, v. 7, n. 4, p. 457-472, 1992.
- GEMAN, S.; GEMAN, D. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 6, n. 6, p. 721-741, 1984.

GOBETTI, S. W.; ORAIR, R. O. Resultado primário e contabilidade criativa: reconstruindo as estatísticas fiscais acima da linha do governo geral. *Texto para Discussão IPEA*, n. 2288, 2017.

HOFFMAN, M. D.; GELMAN, A. The No-U-Turn Sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, v. 15, p. 1593-1623, 2014.

HOLT, C. C. Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting*, v. 20, n. 1, p. 5-10, 2004. (Originalmente ONR Memorandum, 1957).

HYNDMAN, R. J.; ATHANASOPOULOS, G. *Forecasting: principles and practice*. 3. ed. Melbourne: OTexts, 2021.

HYNDMAN, R. J. et al. A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting*, v. 18, n. 3, p. 439-454, 2002.

IMF – INTERNATIONAL MONETARY FUND. *Artificial intelligence in public finance management*. Washington: IMF, 2023. (Fiscal Affairs Department Technical Note).

ISSA, H.; SUN, T.; VASARHELYI, M. A. Research ideas for artificial intelligence in auditing: the formalization of audit and workforce supplementation. *Journal of Emerging Technologies in Accounting*, v. 13, n. 2, p. 1-20, 2016.

KOGAN, A.; SUDIT, E. F.; VASARHELYI, M. A. Continuous online auditing: a program of research. *Journal of Information Systems*, v. 33, n. 1, p. 1-23, 2019.

LAPLACE, P. S. *Théorie analytique des probabilités*. Paris: Courcier, 1812.

LEAL, T. et al. Short-term forecasting of Portuguese GDP using fiscal indicators. *Portuguese Economic Journal*, v. 7, n. 2, p. 73-95, 2008.

MENDES, M. *Rigidez orçamentária no Brasil: causas, consequências e propostas de reforma*. Brasília: IDP, 2022.

MUSGRAVE, R. A. *The theory of public finance: a study in public economy*. New York: McGraw-Hill, 1959.

NEAL, R. M. MCMC using Hamiltonian dynamics. In: BROOKS, S. et al. (ed.). *Handbook of Markov chain Monte Carlo*. Boca Raton: CRC Press, 2011. p. 113-162.

OECD – ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. *Using AI in public finance management*. Paris: OECD Publishing, 2022.

PHUA, C. et al. A comprehensive survey of data mining-based fraud detection research. *Computing Research Repository*, arXiv:1009.6119, 2010.

POISSON, S. D. *Recherches sur la probabilité des jugements en matière criminelle et en matière civile*. Paris: Bachelier, 1837.

TCU – TRIBUNAL DE CONTAS DA UNIÃO. *Relatório de atividades de ciência de dados aplicada à auditoria*. Brasília: TCU, 2021.

WINTERS, P. R. Forecasting sales by exponentially weighted moving averages. *Management Science*, v. 6, n. 3, p. 324-342, 1960.

WIRTZ, B. W.; WEYERER, J. C.; GEYER, C. Artificial intelligence and the public sector: applications and challenges. *International Journal of Public Administration*, v. 42, n. 7, p. 596-615, 2019.

APÊNDICE A – GLOSSÁRIO DE TERMOS TÉCNICOS

Termos Estatísticos e de Inteligência Artificial

Distribuição a posteriori (posterior): Distribuição de probabilidade dos parâmetros de um modelo após a incorporação dos dados observados, combinando a informação prévia (prior) com a verossimilhança dos dados.

Distribuição a priori (prior): Distribuição de probabilidade que representa o conhecimento ou crenças sobre os parâmetros de um modelo antes da observação dos dados. Na abordagem bayesiana, é atualizada pelos dados para produzir a distribuição posterior.

Distribuição de Poisson: Distribuição de probabilidade discreta que expressa a probabilidade de um número de eventos ocorrer em um intervalo fixo de tempo ou espaço, dado que esses eventos ocorrem com taxa média constante e independentemente do tempo desde o último evento.

Hamiltonian Monte Carlo (HMC): Algoritmo de amostragem que utiliza a dinâmica hamiltoniana da física para explorar eficientemente distribuições de probabilidade complexas, produzindo amostras menos correlacionadas que métodos tradicionais de MCMC.

Inferência bayesiana: Método de inferência estatística que utiliza o teorema de Bayes para atualizar a probabilidade de uma hipótese à medida que mais evidências ou informações se tornam disponíveis.

Intervalo de credibilidade: Intervalo derivado da distribuição posterior que contém o parâmetro de interesse com determinada probabilidade. Diferentemente do intervalo de confiança frequentista, permite afirmações diretas sobre a probabilidade do parâmetro estar no intervalo.

MCMC (Markov Chain Monte Carlo): Classe de algoritmos para amostragem de distribuições de probabilidade, construindo uma cadeia de Markov cuja distribuição estacionária é a distribuição desejada. Permite inferência em modelos complexos onde soluções analíticas são intratáveis.

Modelo ARIMA: Modelo Autorregressivo Integrado de Média Móvel. Classe de modelos estatísticos para análise e previsão de séries temporais que combina componentes autorregressivos (dependência de valores passados), diferenciação (para estacionariedade) e média móvel (dependência de erros passados).

Modelo hierárquico: Modelo estatístico com múltiplos níveis, onde os parâmetros de um nível são modelados como provenientes de uma distribuição comum. Permite compartilhamento de informação entre unidades e estimativas mais estáveis para grupos com poucos dados.

R-hat (estatística de Gelman-Rubin): Diagnóstico de convergência para algoritmos MCMC que compara a variância entre cadeias com a variância dentro de cada cadeia. Valores próximos a 1,0 indicam convergência adequada.

Regressão bayesiana: Abordagem para modelagem de regressão que trata os parâmetros como variáveis aleatórias com distribuições de probabilidade, permitindo quantificação natural da incerteza nas estimativas e previsões.

Métricas de Avaliação

AUC (Area Under the Curve): Área sob a curva ROC, medindo a capacidade global de um classificador distinguir entre classes. Valores variam de 0,5 (classificação aleatória) a 1,0 (classificação perfeita).

Brier Score: Métrica que avalia a acurácia de previsões probabilísticas, calculada como a média dos quadrados das diferenças entre as probabilidades previstas e os resultados observados. Valores menores indicam melhor calibração.

Curva ROC (Receiver Operating Characteristic): Representação gráfica do desempenho de um classificador binário, plotando a taxa de verdadeiros positivos contra a taxa de falsos positivos para diferentes limiares de classificação.

Falso positivo: Erro de classificação em que o modelo prevê a ocorrência de um evento que não se materializa. No contexto do SA-BP, corresponde a um alerta emitido para uma situação que não se configurou como evento crítico.

MAE (Mean Absolute Error): Erro Absoluto Médio. Métrica de acurácia que calcula a média das diferenças absolutas entre valores previstos e observados, expressando o erro médio na mesma unidade da variável de interesse.

Recall (Sensibilidade): Proporção de casos positivos reais que são corretamente identificados. No contexto de alertas, mede quantos dos eventos críticos reais foram antecipados pelo sistema.

RMSE (Root Mean Square Error): Raiz do Erro Quadrático Médio. Métrica de acurácia que calcula a raiz quadrada da média dos quadrados dos erros, penalizando mais fortemente erros grandes.

Taxa de detecção: Proporção de eventos críticos que foram corretamente identificados pelo sistema de alerta dentro do horizonte de antecipação definido. Equivalente ao recall no contexto de sistemas de alerta precoce.