



ANÁLISE DE SOBREVIVÊNCIA PARA A AVALIAÇÃO DOS ASPECTOS INFLUENTES NO TEMPO DE DESEMPREGO

Allan Rafael Dias Araújo Severino¹, <https://orcid.org/0009-0007-7894-1408>

Denilson de Oliveira Silva², <https://orcid.org/0009-0000-4031-7772>

¹Departamento de Estatística, Universidade Estadual da Paraíba, Campina Grande, Brasil.

Email: allan.severino@aluno.uepb.edu.br;

²Departamento de Estatística, Universidade Estadual da Paraíba, Campina Grande, Brasil.

Email: denilson.oliveira.silva@aluno.uepb.edu.br;

RESEARCH ARTICLE

Annals of Data Analysis
<https://periodicos.ufpe.br/revistas/ADA/>
ISSN Online: xxxx-xxxx

Submitted on: 12.12.2025.

Accepted on: 03.08.2026.

Published on: 05.08.2026.

Copyright © 2026 by author(s).

This work is licensed under the Creative Commons Attribution International License CC BY-NC-ND 4.0
<http://creativecommons.org/licenses/by-nc-nd/4.0/deed.en>



Abstract

Unemployment is a central issue for the labor market and public policy. This study analyzed the behavior of individuals in this situation, using survival analysis to estimate the time until the first formal job. The results showed that individuals who participated in unemployment insurance programs increased the likelihood of remaining unemployed for longer. Age also plays a role, particularly among those aged 50 to 61, while the 30 to 39 age group showed no significant effect. The conclusion is that unemployment results from the interaction between individual characteristics and economic factors, including access to social support programs.

Keywords

Keywords: Unemployment; Survival Analysis; Unemployment Insurance; Labor Market.

1. Introdução

O desemprego configura-se como uma condição em que indivíduos, motivados pela necessidade de trabalho, enfrentam obstáculos para alcançar uma ocupação, mesmo que estejam disponíveis e dispostos a trabalhar. Suas causas são múltiplas e complexas: podem derivar de choques conjunturais, como crises econômicas, pandemias ou recessões; de mudanças estruturais no mercado de trabalho, como automação, desindustrialização ou o avanço tecnológico que substitui a mão de obra; ou ainda de fatores institucionais e

regulatórios, como a rigidez da legislação trabalhista ou a insuficiente oferta de qualificação profissional (OCDE, 2020; OIT, 2024).

Esse fenômeno vai além do âmbito individual, gerando impactos macroeconômicos expressivos. O desemprego reduz o consumo das famílias, enfraquece a demanda agregada e pode frear o crescimento econômico. Em termos sociais, amplia as desigualdades, pois grupos historicamente vulneráveis — pessoas com menor escolaridade, mulheres, idosos e moradores de regiões menos desenvolvidas — enfrentam maiores dificuldades de reinserção e períodos mais longos sem ocupação (World Bank, 2022; OIT, 2024). A persistência dessa condição compromete a segurança financeira, acentua a pobreza e pode gerar efeitos psicológicos e sociais duradouros, incluindo a perda de autoestima e impactos sobre a saúde mental (SEN, 1999; STIGLITZ, 2019).

Diante desse cenário, compreender os fatores determinantes da duração do desemprego é essencial para o planejamento de políticas públicas e estratégias de inclusão produtiva. Nesse contexto, a análise de sobrevivência surge como uma ferramenta estatística adequada para estudar o tempo até a ocorrência de um evento específico, como a recolocação profissional. Essa metodologia permite avaliar o risco de saída do desemprego ao longo do tempo, considerando variáveis socioeconômicas, demográficas e institucionais que podem afetar esse processo (KLEIN & MOESCHBERGER, 2003; HOSMER et al., 2008).

Neste estudo, propõe-se analisar a duração do desemprego entre indivíduos economicamente ativos, investigando como fatores socioeconômicos (salário, participação em programas de assistência), demográficos (faixa etária) e institucionais influenciam o tempo até a obtenção de um novo emprego formal. Busca-se não apenas estimar a influência dessas variáveis sobre o tempo de desemprego, mas também discutir implicações práticas para políticas públicas, contribuindo para a formulação de medidas mais eficazes na mitigação das desigualdades laborais e na promoção da inclusão econômica.

Ao aplicar uma abordagem quantitativa rigorosa, fundamentada em técnicas de modelagem estatística, o trabalho pretende aprofundar o entendimento sobre os determinantes da permanência no desemprego, fornecendo evidências empíricas que podem subsidiar a tomada de decisão em nível governamental e institucional.

2. Objetivo

Analisar os fatores que influenciam a duração do desemprego e a probabilidade de reintegração ao mercado de trabalho, considerando o impacto do seguro-desemprego, da expectativa salarial e da idade.

Objetivos específicos

- identificar que fatores aumentam ou diminuem o tempo que uma pessoa permanece desempregada
- Verificar se o recebimento de seguro-desemprego, e a idade afetam a chance de alguém conseguir um emprego

- A probabilidade de uma pessoa ainda estar desempregada após um certo período de tempo
- Através da curva de sobrevivência de Kaplan-Meier, visualizar como o tempo de desemprego varia para diferentes grupos, comparando o tempo de desemprego entre pessoas que recebem seguro-desemprego e aquelas que não recebem, ou como a idade e o salário, impactam a duração do desemprego.

3. Materias e Métodos

Como o intuito deste estudo é avaliar os fatores que influenciam no desemprego. Para isso, utilizaremos um banco de dados reais obtidos no kaggle¹, contendo 3343 observações e 5 variáveis. Com as seguintes variáveis idade, seguro desemprego, salário, Período de desemprego, e o evento se o indivíduo conseguiu ou não emprego que é a variável resposta. A seguir, na Tabela 1, apresentamos uma descrição detalhada das variáveis que compõem o nosso banco de dados. Cada variável é especificada quanto ao seu tipo, unidade de medida e possíveis categorias, permitindo uma compreensão clara dos dados utilizados na análise.

Tabela 1: Descrição das variáveis da base de dados.

Variável	Descrição
Período de desemprego	Refere-se ao número de períodos que uma pessoa permanece desempregada.
Evento de interesse	Representa se uma pessoa encontrou um emprego (1) ou não (0).
Seguro desemprego	Indica se o indivíduo recebe seguro desemprego.
Salário	Refere-se à transformação logarítmica do salário
Idade	Indica a idade do indivíduo

Fonte: Elaborado pelos autores, (2025).

3.1. Análise de sobrevivência.

De acordo com a literatura estatística, a análise de sobrevivência é uma técnica poderosa usada para estudar o tempo até a ocorrência de um evento específico, como a obtenção de emprego, a recuperação em tratamentos de saúde ou a falência de empresas. Essa abordagem é amplamente aplicada em diversas áreas, como medicina, economia e ciências sociais, onde o interesse está em compreender como variáveis preditoras influenciam a "sobrevivência" até o evento de interesse (Colosimo & Giolo, 2006; Klein & Moeschberger, 2003). Também temos a falha é considerada o evento de interesse em nossa análise, representando o desfecho que buscamos observar ao longo do período treinado. Em nosso trabalho, esse evento de interesse é definido pelo fato de uma pessoa conseguir ou não um emprego. Assim, o estudo se concentra em identificar os fatores associados à ocorrência desse evento, proporcionando uma melhor compreensão dos aspectos que influenciam a transição do desemprego para o emprego. Dado por:

$$\delta \begin{cases} 0, & \text{Censura} \\ 1, & \text{Falha} \end{cases}$$

3.2. Função de sobrevivência de Kaplan Meier.

¹ <https://www.kaggle.com/datasets/kunduruani/survival-unemployment>

A análise de sobrevivência é um conjunto de técnicas estatísticas utilizadas para estudar dados cujo objetivo principal é avaliar o tempo até a ocorrência de um evento específico. Esses eventos podem variar, incluindo falecimento de um paciente, falha de um componente, recuperação de uma doença, entre outros. Neste trabalho, será utilizado o estimador da função de sobrevivência de Kaplan-Meier, proposta por Kaplan e Meier em 1958. A expressão matemática do estimador é uma função degrau, $\hat{S}(t)$, que apresenta saltos nos tempos observados de falha é a seguinte:

$$\hat{S}(t) = \frac{n \text{ de observações que não falharam até o tempo } t}{n \text{ total de observações}}$$

3.3. Função densidade de sobrevivência.

A função de densidade de sobrevivência descreve a taxa de ocorrência de falhas ou eventos ao longo do tempo em uma análise de sobrevivência. É uma medida que indica a probabilidade instantânea de que o evento ocorra em um pequeno intervalo de tempo após um determinado ponto, dado que o indivíduo, ou item ainda não ocorreu o evento. Geralmente representado como $\hat{f}(t)$, ela está relacionada com a função de sobrevivência $\hat{S}(t)$, pois descreve o tempo exato de falha ou evento. A densidade de sobrevivência é útil para entender a distribuição temporal dos eventos e calcular outras métricas, como a função de risco, (ALMEIDA, 2016).

3.4. Estimador de Kaplan Meier.

O estimado de Kaplan-Meier, desenvolvido por Edward L. Kaplan e Paul Meier em 1958, é uma das ferramentas mais utilizadas na análise de sobrevivência. Ele permite estimar a função de sobrevivência de uma população com base em uma amostra, especialmente quando há censuras (indivíduos que não tiveram o evento de interesse observado até o fim do estudo). A função de sobrevivência $\hat{S}(t)$, estimado por Kaplan-Meier, é uma função em degraus, onde cada degrau ocorre em um tempo de falha observada, sendo ajustado com base na quantidade de indivíduos em risco naquele instante. Para calcular o estimador, considera-se a probabilidade de sobrevivência acumulada até cada ponto de falha, multiplicando-se como probabilidades de sobrevivência de cada intervalo, (ZHOU, 2025). O método é amplamente utilizado em medicina, engenharia e outras áreas para compreender o tempo até eventos como falhas ou mortes, dada por:

$$\hat{S}(t) = \prod_{j:t_j < t} \left(\frac{n_j - d_j}{n_j} \right) = \prod_{j:t_j < t} \left(1 - \frac{d_j}{n_j} \right)$$

Em que:

t: tempos;

n_j : número de indivíduos em risco;

d_j : número de falhas.

3.5. Teste log rank e wilcoxon

O teste Log-Rank é uma ferramenta não paramétrica usada para comparar curvas de sobrevivência de dois ou mais grupos, focando em diferenças no tempo até a ocorrência de um evento. Ele assume que as curvas de sobrevivência devem ter diferenças constantes ao longo do tempo (Klein & Moeschberger, 2003; Collett, 2015). Já o teste de Wilcoxon, também compara curvas de sobrevivência, mas é mais sensível às diferenças iniciais entre as curvas, sendo influenciado principalmente pelos eventos que ocorrem no início do período de acompanhamento (Lee & Wang, 2013). Ambos os testes são amplamente utilizados em estudos médicos e experimentais para verificar se as funções de sobrevivência são significativamente diferentes entre grupos, com base em hipóteses que compartilham diferentes padrões de variação entre as curvas (Hosmer, Lemeshow & May, 2008). E as suas hipóteses são:

$$\begin{cases} H_0: \mu_1 = \mu_2 \\ H_1: \mu_1 \neq \mu_2 \end{cases}$$

3.6. Modelo de riscos proporcionais de Cox (semi-paramétrico)

O modelo de riscos proporcional de Cox semi paramétrico, também conhecido como modelo de riscos proporcionais de Cox, é uma ferramenta popular na análise de sobrevivência. Proposto por David Cox em 1972, ele permite estimar o efeito de variáveis explicativas sobre o tempo até a ocorrência de um evento sem assumir uma distribuição específica para os tempos de falha. O modelo expressa a taxa de risco como uma função exponencial das covariáveis, assumindo que os riscos são proporcionais ao longo do tempo para diferentes grupos. Essa flexibilidade torna o modelo de Cox ideal para estudos onde se deseja avaliar o impacto de variáveis como idade, tratamento ou características demográficas sobre o risco de um evento. Dado por:

$$\lambda(t|x) = \lambda_0(t) \exp(x_1\beta_1 + x_2\beta_2 + \dots + x_p\beta_p) = \lambda_0(t) \exp(X\beta)$$

Em que:

$\lambda(t|x)$: representa a taxa de risco condicional no tempo(t), dado o vetor de covariáveis(X);

$\lambda_0(t)$: é a taxa de risco de base no tempo(t);

$\exp(X\beta)$: é a notação vetorial compacta para $(x_1\beta_1 + x_2\beta_2 + \dots + x_p\beta_p)$.

3.7. método stepwise

O método *stepwise* é uma técnica de seleção de variáveis bastante utilizada em análises de regressão. Ele funciona por meio da inclusão ou exclusão de variáveis em etapas sucessivas, permitindo identificar o modelo mais adequado com base nos dados disponíveis. Existem três abordagens principais desse método, a seleção progressiva (*forward stepwise*), a seleção regressiva (*backward stepwise*) e uma abordagem mista que combina as duas (*both stepwise*). (DRAPER & SMITH, 1998).

3.8 Métricas Utilizadas na Análise de Sobrevivência

➤ **Concordância**

A métrica de concordância, também conhecida como *C-index*, avalia a capacidade preditiva do modelo de sobrevivência, indicando o quão bem as previsões se alinham com os tempos observados. Valores próximos de 1 indicam excelente desempenho, enquanto valores próximos de 0,5 representam desempenho aleatório (HARRELL et al., 1996).

➤ **Teste Log-rank**

O teste Log-rank é amplamente utilizado para comparar curvas de sobrevivência entre dois ou mais grupos. Ele verifica se há diferenças estatisticamente significativas nas probabilidades de sobrevivência ao longo do tempo (KLEIN; MOESCHBERGER, 2003).

➤ **Máxima Verossimilhança**

O método da máxima verossimilhança é utilizado para estimar os parâmetros do modelo de sobrevivência, buscando os valores que maximizam a probabilidade de observar os dados amostrais. Essa abordagem garante estimativas consistentes e eficientes (COLOSIMO; GIOLO, 2006).

➤ **Teste de Wald**

O teste de Wald avalia a significância individual dos coeficientes no modelo de regressão de sobrevivência, verificando se cada parâmetro difere estatisticamente de zero. Valores de p inferiores a 0,05 indicam associação significativa entre a variável e o tempo até o evento (HOSMER; LEMESHOW; MAY, 2008).

4. Resultados e discussões

Inicialmente, de acordo com a análise descritiva das variáveis, podemos analisar alguns pontos interessantes. Os dados da tabela 2, mostram um grupo com idade variando entre 20 e 61 anos, com idade média de 35 anos e mediana de 34 anos, visando uma distribuição etária relativamente jovem, equipamentos em torno dos 30 - 40 anos. Em termos de salário, os valores variam de R\$ 230,27 a R\$ 754,59, com uma média de R\$ 533,61 e mediana de R\$ 562,75, fazendo com que a maioria dos movimentos esteja entre esses valores centrais. O período de desemprego varia extensamente, de 1 a 28 meses, com uma média de 6,35 meses, mas a mediana de 5 meses sugere que metade das pessoas teve um tempo de desemprego até esse limite. O primeiro quartil indica que 25% dos indivíduos tiveram até 0 meses de desemprego, enquanto 75% tiveram até 1 mês.

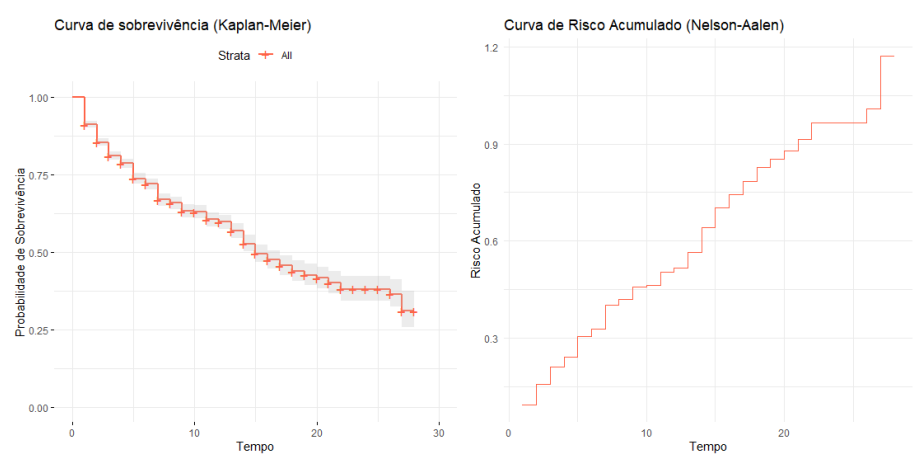
Tabela 2: Medidas De Posição.

Medidas	Idade	Salário	Período de desemprego
Min	20,00	230,27	1,00
Máx	61,00	754,59	28,00
1º Quartil	27,00	524,42	0,00
3º Quartil	43,00	601,250	1,00
Média	35,00	533,61	6,35

Mediana	34,00	562,752	5,00
Fonte: Elaborado pelos Autores, (2025).			

Na figura 1, a curva Kaplan-Meier mostra a probabilidade de uma pessoa permanecer desempregada ao longo do tempo. Inicialmente, todos estão desempregados, mas essa chance diminui conforme algumas pessoas encontram emprego. Cada queda na curva representa alguém que conseguiu trabalhar, enquanto o "número em risco" abaixo indica quantos ainda estão desempregados em momentos específicos. A curva revela que o desemprego diminuiu ao longo do tempo, embora algumas pessoas permaneçam desempregadas até ao final da observação. Esse comportamento é coerente com o que se observa na literatura, uma vez que a análise de sobrevivência permite estimar a permanência em determinada condição ao longo do tempo e identificar grupos que permanecem em risco por períodos mais longos (BUSTAMANTE-TEIXEIRA; FAERSTEIN; LA TORRE, 2002).

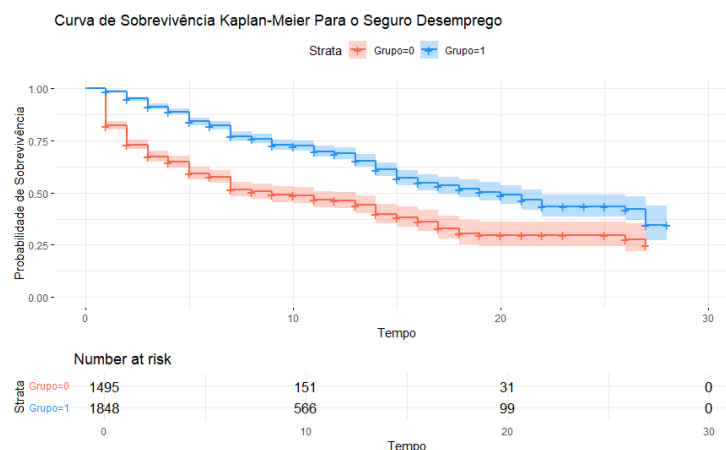
Figura 1: Curva de kaplan meier para o período de desemprego.



Fonte: Elaborado pelos autores, (2025).

Na figura 2 tem-se a curva Kaplan-Meier apresentada mostra a probabilidade de permanência no seguro-desemprego ao longo do tempo, comparando dois grupos: aqueles que não recebem (grupo 0) e aqueles que recebem (grupo 1). Observa-se que, inicialmente, ambos os grupos começam com a mesma probabilidade de sobrevivência. Com o passar do tempo, o grupo que não recebe (linha vermelha) apresenta uma redução mais acentuada na probabilidade de permanência no seguro-desemprego em comparação ao grupo que recebe (linha azul). Ao final do período analisado, a probabilidade de não ter seguro-desemprego é visivelmente menor para o grupo que não recebe, diminuindo que aqueles que recebem o benefício tendem a permanecer mais tempo nele. Além disso, os números abaixo dos gráficos, o número de risco indica quantas pessoas de cada grupo permaneceram no estudo ao longo do tempo.

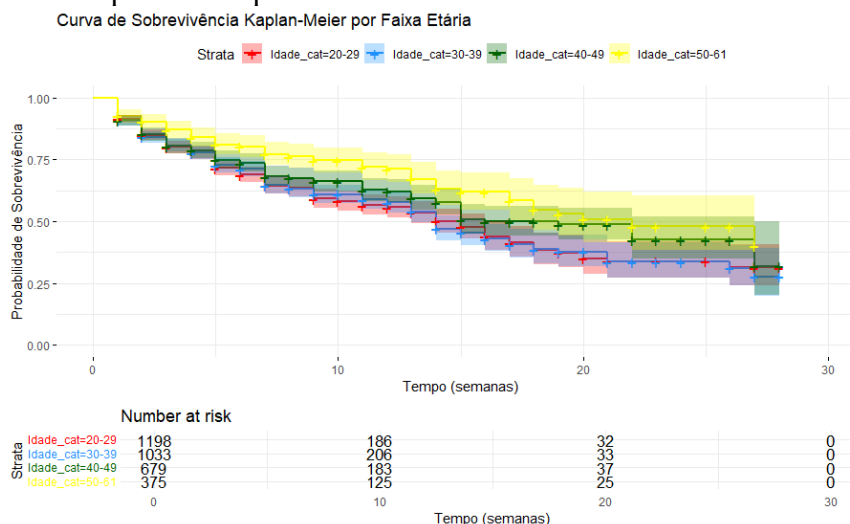
Figura 2: Curva de kaplan meier para o seguro desemprego.



Fonte: Elaborado pelos Autores, (2025).

Na figura 3, a curva Kaplan-Meier apresentada analisa a probabilidade das ao analisar as curvas por faixa etária, observamos que os trabalhadores mais jovens (20-29 anos e 30-39 anos) tendem a sair do desemprego mais rapidamente, pois suas probabilidades de permanecerem desempregados diminuem de forma mais acentuada ao longo dos períodos. Por outro lado, os trabalhadores mais velhos (40-49 anos e especialmente 50-61 anos) apresentam uma probabilidade de desemprego mais alta e estável, diminuindo que esses grupos permanecem desempregados por períodos mais longos. Isso sugere que a idade influencia a duração do desemprego, com os trabalhadores mais velhos possivelmente enfrentando maiores dificuldades para encontrar novas oportunidades de trabalho rapidamente, enquanto os trabalhadores mais jovens conseguem se reinserir no mercado com maior facilidade ao longo do tempo.

Figura 3: Curva de kaplan meier para a idade.

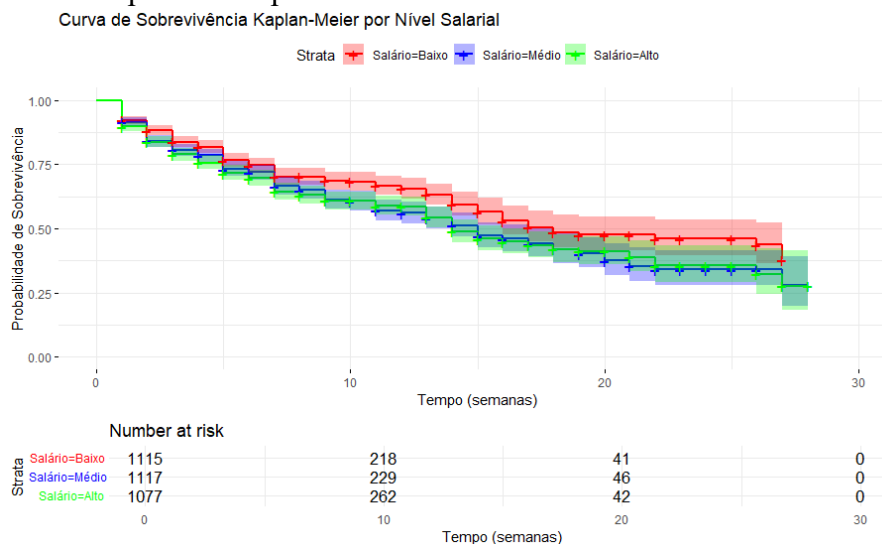


Fonte: Elaborado pelos Autores, (2025).

Na figura 4 podemos observar que os três grupos começam com probabilidades de desemprego semelhantes. No entanto, ao longo do tempo, o grupo de Salário-Baixo apresenta uma redução mais rápida na probabilidade de permanência desempregada, enquanto os grupos de Salário-Médio e Salário-Alto apresentam curvas mais próximas entre si. Isso pode indicar que pessoas com períodos mais baixos têm uma probabilidade um pouco maior de sair do

desemprego mais cedo, enquanto aquelas com períodos médios e altos tendem a permanecer desempregadas por períodos mais longos, embora as diferenças entre os três grupos sejam relativamente sutis. Os número em risco, mostram a quantidade de pessoas de cada nível salarial ainda desempregadas ao longo do tempo, com uma queda acentuada na quantidade de indivíduos nos períodos finais.

Figura 4: Curva de kaplan meier para o Salário.



Fonte: Elaborado pelos Autores, (2025).

Os resultados da Tabela 3, com $p < 0,001$ em ambos os testes log-rank e Wilcoxon para seguro-desemprego, idade e salário, encontram respaldo em estudos recentes. Por exemplo, Adinkra-Darko (2024) evidenciou que escolaridade, idade e gênero influenciam significativamente a duração do desemprego em Gana. Pathiranage et al. (2025) mostram vulnerabilidade elevada para jovens e residentes rurais com baixa escolaridade, reforçando que características individuais têm papel central na saída do desemprego. No contexto brasileiro, análises recentes sobre informalidade também apontam que trabalhadores com menor escolaridade enfrentam períodos mais longos de desemprego.

Tabela 3: Teste log-rank e wilcoxon.

	Wilcoxon	Log - rank
Seguro desemprego	<0,001	<0,001
Idade	<0,001	<0,001
Salário	<0,001	<0,001

Fonte: Elaborado pelos Autores, (2025).

A tabela 4, tem-se o modelo inicial apresentada abaixo e nele examinamos as relações entre as variáveis independente e a variável dependente (Período de desemprego e evento). A tabela resume os coeficientes de regressão estimados para cada variável independente, juntamente com outras estatísticas relevantes, como p-valor, e o valor z. No modelo inicial apenas duas das variáveis não foi significativa, ou seja, não contribuiu de forma relevante para o desemprego. Por isso, precisamos aplicar um método de seleção de variáveis para identificar quais variáveis são realmente importantes e quais devem ser removidas. A seleção de variáveis é importante porque melhora a precisão e a interpretabilidade do modelo, removendo variáveis

irrelevantes. Aplicando esses métodos, podemos comparar diferentes modelos e escolher o que apresenta o melhor desempenho com o menor número de variáveis. Assim, garantimos que o modelo final seja eficiente e fácil de interpretar.

Tabela 4: Modelo de cox semi-paramétrico inicial.

	Coefficiente	Exp (Coeficiente)	Erro padrão (Coeficiente)	p - valor
Salário Médio	0,441	1,556	0,081	<0,001 ***
Salário baixo	0,556	1,743	0,084	<0,001 ***
Idade 30-39	0,055	1,056	0,076	0,471
Idade 40-49	-0,084	0,918	0,088	0,340
Idade 50-61	-0,309	0,735	0,116	0,007 **
Seguro Desemprego	-1,026	0,358	0,066	<0,001 ***

Fonte: Elaborado pelos autores, (2025).

Na Tabela 5, temos o processo de seleção de variáveis utilizando o método Stepwise, todas logo após a aplicação todas as variáveis do modelo inicial foram mantidas no modelo final. Isso indica que, de acordo com os critérios de seleção adotados, todas as variáveis envolvidas contribuem para a explicação das características do estudo. Portanto, o modelo final inclui todas as variações originalmente propostas, indicando que cada uma delas possui alguma relevância para a previsão dos períodos de desemprego e que a remoção de qualquer uma poderia comprometer a capacidade explicativa do modelo.

Com isso temos que para o Salário Médio com um coeficiente de 0,441 indicando, que as pessoas nessa faixa tendem a permanecer mais tempo desempregadas. Para Salário Baixo o coeficiente foi de 0,556, a probabilidade é ainda maior, indicando menos oportunidades de arrumar emprego. Na faixa 30-39 anos temos 0,055. então a influência é mínima. Para 40-49 anos (-0,084), há uma leve redução no desemprego, enquanto para 50-61 anos (-0,309) essa redução no desemprego se torna ainda maior, possivelmente devido à experiência. E por último o Seguro Desemprego com (-1.026) associa-se a uma redução significativa na probabilidade de desemprego, redução do incentivo ao reemprego ou suporte temporário.

Tabela 5 : Modelo final de cox semi-paramétrico do stepwise.

	Coefficiente	Exp (Coeficiente)	Erro padrão (Coeficiente)	p - valor
Salário Médio	0,441	1,556	0,081	<0,001 ***
Salário baixo	0,556	1,743	0,084	<0,001 ***
Idade 30-39	0,055	1,056	0,076	0,471
Idade 40-49	-0,084	0,918	0,088	0,340
Idade 50-61	-0,309	0,735	0,116	0,007 **
Seguro Desemprego	1,026	0,358	0,066	<0,001 ***

Fonte: Elaborado pelos autores, (2025).

Os resultados da Tabela 6, indicam que o modelo final é bem ajustado e estatisticamente significativo e bem fundamentado. A concordância de 0,79 mostra que o modelo tem desempenho um desempenho bom em prever a ordem dos tempos de sobrevivência. Os testes de Logrank, máxima verossimilhança e Wald, todos com p - valor < 0,001, indicam que as variáveis explicativas têm impacto relevante no resultado e que o modelo é significativamente

melhor do que um modelo sem variáveis. Isso reforça a adequação do modelo aos dados analisados.

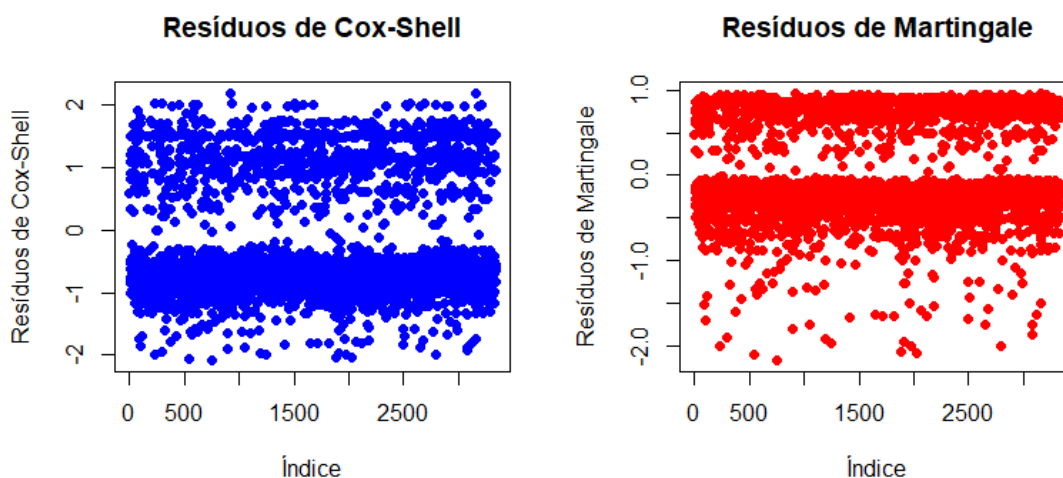
Tabela 6: Métricas do modelo final.

Métricas	Estatísticas
Concordância	0,79
Teste Logrank	<0,001
máxima verossimilhança	<0,001
Teste wald	<0,001

Fonte: Elaborado pelos autores, (2025).

A Figura 5, tem-se a análise gráfica dos resíduos mostra que eles estão distribuídos de forma desejada, sem qualquer organização ou padrão visível. No gráfico de resíduos de Cox-Shell, os pontos estão espalhados ao redor do eixo zero, indicando que o modelo ajustado funciona bem para os dados, sem apresentar problemas. Nos gráficos de resíduos de Martingale, os pontos também seguem uma distribuição sem ordem ou padrão, o que confirma que o modelo respeita os requisitos necessários. De forma geral, esses resultados mostram que o modelo é adequado e capturou bem as relações entre as variáveis, sem deixar erros importantes.

Figura 5: Análise de Resíduos do modelo final.



Fonte: Elaborado pelos autores, (2025).

A Tabela 6, apresenta os resultados do teste de Schoenfeld para avaliar a suposição de proporcionalidade dos riscos no modelo de regressão de sobrevivência. Observa-se que nenhuma das variáveis analisadas violou significativamente a suposição de proporcionalidade, considerando que os valores de p foram todos maiores que 0,05, Salário ($p = 0,18$), Idade ($p = 0,15$) e Seguro Desemprego ($p = 0,09$). O teste global apresentou $p = 0,08$, indicando que, em conjunto, as variáveis não apresentam evidências de violação da suposição de riscos proporcionais.

Tabela 6: Teste de schoenfeld.

Variáveis	p - valor
Salário	0,18
Idade	0,15
Seguro desemprego	0,09
Global	0,08

Fonte: Elaborado pelos autores, (2025).

5. Conclusão

Com base nos resultados obtidos, o modelo demonstra que fatores socioeconômicos e demográficos exercem influência significativa sobre a probabilidade de desemprego. Observou-se que indivíduos com salários médios ou baixos e aqueles que participam de programas de seguro-desemprego apresentam maior propensão a permanecer desempregados por períodos mais longos, o que sugere uma relação direta entre vulnerabilidade econômica e dificuldade de reinserção no mercado de trabalho.

Além disso, a faixa etária de 50 a 61 anos mostrou-se estatisticamente significativa, indicando que o avanço da idade pode representar uma barreira à recolocação profissional. Esse resultado reforça a necessidade de políticas específicas voltadas à inclusão de trabalhadores mais velhos, que frequentemente enfrentam discriminação etária e menor acesso a oportunidades de capacitação. Por outro lado, indivíduos entre 30 e 39 anos não apresentaram impacto relevante sobre o tempo de desemprego, o que pode estar associado à maior flexibilidade e adaptação desse grupo etário ao mercado.

Esses achados revelam que o desemprego é um fenômeno complexo, influenciado tanto por características individuais quanto por fatores estruturais e institucionais, como o próprio funcionamento dos programas de seguro-desemprego. Dessa forma, as políticas públicas devem ser desenhadas considerando essas múltiplas dimensões, de modo a reduzir desigualdades e fortalecer a empregabilidade. Portanto, o estudo contribui ao evidenciar que o enfrentamento do desemprego requer ações integradas, que combinem incentivos à qualificação profissional, estímulos à contratação de grupos vulneráveis e revisão das condições de acesso e permanência em programas assistenciais. A incorporação dessas evidências pode orientar gestores públicos e pesquisadores na formulação de estratégias mais eficientes e socialmente justas para promover a inclusão econômica e o desenvolvimento sustentável.

Acknowledgements

Os autores reconhecem a relevância do WANDA 2025 como fórum de disseminação científica, cuja organização e estrutura possibilitaram a apresentação e o aprimoramento dos resultados desta pesquisa.

Conflict of Interest Declaration

Os autores declaram que este trabalho foi desenvolvido de forma independente, sem a existência de conflitos de interesse. Todos os coautores aprovaram integralmente o conteúdo do manuscrito e não possuem vínculos financeiros relacionados ao estudo.

6. Referências

- Organisation for Economic Co-operation and Development. (2020). *COVID-19 and international trade: Issues and actions*. OECD Publishing.
- International Labour Organization. (2024). *World employment and social outlook: Trends 2024*. ILO.
- World Bank. (2022). The economic impacts of the COVID-19 crisis. In *World development report 2022*. World Bank.
- Colosimo, E. A., & Giolo, S. R. (2021). *Análise de sobrevivência aplicada*. Editora Blucher.
- Koenker, R., & Bassett, G. (1979). Coefficient variation. *Econometrica*, 47(5), 1287–1294.
- Organisation for Economic Co-operation and Development. (2020). *Employment outlook 2020: Worker security and the COVID-19 crisis*. OECD Publishing.
- Lee, E. T., & Wang, J. W. (2013). *Statistical methods for survival data analysis* (4th ed.). John Wiley & Sons.
- Hosmer, D. W., Lemeshow, S., & May, S. (2008). *Applied survival analysis: Regression modeling of time-to-event data* (2nd ed.). John Wiley & Sons.
- Klein, J. P., & Moeschberger, M. L. (2003). *Survival analysis: Techniques for censored and truncated data* (2nd ed.). Springer.
- Kunduru, A. (n.d.). *Survival unemployment* [Data set]. Kaggle. <https://www.kaggle.com/datasets/kunduruanil/survival-unemployment>
- Bustamante-Teixeira, M. T., Faerstein, E., & La Torre, M. R. (2002). Técnicas de análise de sobrevida. *Cadernos de Saúde Pública*, 18(3), 579–594. <https://doi.org/10.1590/S0102-311X2002000300008>
- Adinkra-Darko. (2024). *Unemployment duration and its covariates*.
- Pathiranage, et al. (2025). *Integrating survival analysis and machine learning techniques for modeling and predicting unemployment duration*.
- Almeida, T. A. C. (2016). *Análise de sobrevivência: Conceitos e aplicações*. Universidade de Brasília.
- Zhou, Y. (2025). Wasserstein–Kaplan–Meier survival regression. *Journal of Statistical Computation and Simulation*.
- Draper, N. R., & Smith, H. (1998). *Applied regression analysis* (3rd ed.). Wiley.
- Colosimo, E. A., & Giolo, S. R. (2006). *Análise de sobrevivência aplicada*. Edgard Blücher.
- Harrell, F. E., Lee, K. L., Matchar, D. B., & Reichert, T. A. (1996). Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine*, 15(4), 361–387.
- Hosmer, D. W., Lemeshow, S., & May, S. (2008). *Applied survival analysis: Regression modeling of time-to-event data* (2nd ed.). Wiley.
- Klein, J. P., & Moeschberger, M. L. (2003). *Survival analysis: Techniques for censored and truncated data* (2nd ed.). Springer.

