

**GUARDA AMOSTRAL PARA PROCESSOS QUE SERÃO ELIMINADOS  
TRIBUNAL REGIONAL DO TRABALHO DA 7ª REGIÃO  
SECRETARIA DE GOVERNANÇA E GESTÃO ESTRATÉGICA  
COORDENADORIA DE ESTATÍSTICA**

## **1. Introdução**

Neste relatório serão apresentados a metodologia e os resultados da aplicação da técnica estatística chamada Amostragem Estratificada Proporcional, utilizada para criação de uma guarda amostral para processos que serão eliminados pelo TRT7.

Proposta pelo Conselho Nacional de Justiça ([https://www.cnj.jus.br/wp-content/uploads/2021/02/Manual\\_de\\_Gestao\\_Documental.pdf](https://www.cnj.jus.br/wp-content/uploads/2021/02/Manual_de_Gestao_Documental.pdf)), a aplicação desta técnica, tem por objetivo permitir ao usuário do acervo preservado reconstruir procedimentos, técnicas e normas utilizadas à época da criação do documento, bem como dar condições ao usuário de fazer pesquisas por meio da documentação preservada. Além disso, a guarda amostral também serve para representar as funções e atividades do órgão em determinado período.

## **2. Metodologia**

### **i. Plano Amostral**

Para este plano amostral, desenvolvido para o armazenamento de um acervo documental, não se sabe ao certo o objeto das pesquisas que serão realizadas no futuro, ou seja, o parâmetro que se quer estimar, um dos elementos mais importantes na seleção da amostra, neste caso não é conhecido. Mediante a esta falta de informação, a proposta se baseia na seleção de amostras para estimar proporções, adotando valores que gerem maior variabilidade da estimativa em estudo possível, o que irá gerar uma amostra baseada em critério mais conservador.

Há dois tipos de amostragem: as não probabilísticas e as probabilísticas. As amostras feitas por julgamento do pesquisador são não probabilísticas. Na amostragem probabilística, a probabilidade de seleção de cada item ou indivíduo da população é conhecida, fazendo que seja possível estimar o nível de erro. Além disso, neste tipo é possível generalizar o resultado para toda a população.

No presente estudo será utilizada a técnica de Amostragem Estratificada Proporcional. Esta é realizada na presença de muita heterogeneidade dos dados da população. A ideia é dividir a população em partições (estratos) de forma a proporcionar indivíduos homogêneos dentro dos grupos e heterogêneos entre os grupos.

O CNJ sugere que as amostras sejam estratificadas por ano de distribuição ou por assuntos repetitivos, desta forma, para este presente plano amostral, considerou-se como subpopulações os anos de distribuição das ações.

## ii. Cálculo do tamanho da amostra

Para o cálculo do tamanho de uma amostra estratificada proporcional, quando posteriormente será utilizada uma amostragem aleatória simples sem reposição dentro dos estratos, aplica-se a seguinte fórmula matemática:

$$n = \frac{\sum_{h=1}^H N_h^2 p_h (1 - p_h) / W_h}{N^2 D + \sum_{h=1}^H N_h p_h (1 - p_h)}$$

em que,

- $p_h$  = proporção de indivíduos com a característica em estudo no estrato  $h$ ;
- $H$  = número de estratos;
- $W_h = \frac{N_h}{N}$ ;
- $D = \frac{B^2}{Z_\alpha^2}$ , sendo  $B$  o erro máximo desejado e  $Z_\alpha$  o quantil da distribuição normal padrão, em que  $1 - \alpha$  corresponde ao nível de confiança da amostragem;
- $N$  = tamanho da população;
- $N_h$  = tamanho da população no  $h$ -ésimo estrato.

Este estudo está sendo considerado para o caso em que se deseja estimar a proporção de uma população. Logo, a variável em análise segue uma distribuição *bernoulli* e tem média dada por  $E(x) = p$  e variância dada por  $Var(x) = p(1 - p)$ , sendo  $p$  a proporção de indivíduos na população com a característica em análise. Desta forma, não sabendo qual o interesse do pesquisador na futura amostra, indica-se utilizar  $p = 0,5$ , garantindo a maior variância possível.

Seguindo as sugestões do CNJ, utilizou-se no presente estudo, um Erro máximo ( $B$ ) de 0,03 e um nível de confiança de 97% que gera um valor de  $Z_\alpha = 2,17$ . Portanto, considerando  $p = 0,5$  e as sugestões do erro máximo permitido e do nível de confiança adotado, pode-se calcular o valor de  $n$  através de:

$$n = \frac{0,25N}{N \left( \frac{B^2}{Z_\alpha^2} \right) + 0,25}$$

Por fim, após a determinação do tamanho da amostra, utilizou-se a alocação proporcional para encontrar a amostra correspondente a cada estrato ( $n_h$ ). Este cálculo, consiste na distribuição proporcional do tamanho da amostra em relação ao tamanho do estrato.

$$n_h = nW_h = n \frac{N_h}{N}$$

### 3. Resultados

A Tabela 1 apresenta a relação de estratos e suas respectivas quantidades utilizadas no cálculo do plano amostral.

*Tabela 1: Relação de estratos e suas respectivas quantidades*

ANO	QUANTIDADE
1987	3
1989	16
1990	3
1991	8
1992	52
1993	291
1994	1728
1995	1491
1996	1331
1997	1427
1998	609
1999	904
2000	15
2001	3
2005	2
TOTAL	7883

Fonte: Tribunal Regional do Trabalho da 7ª Região

Utilizando as quantidades apresentadas acima, calculou-se o tamanho da amostra como sugerido pelo CNJ, utilizando  $B = 0,03$  e  $1 - \alpha = 97\%$ , obtendo um tamanho amostral de 1.131. Este valor foi alocado proporcionalmente a cada estrato de acordo com a alocação proporcional e seus resultados são encontrados na Tabela 2.

Tabela 2: Amostra alocada a cada estrato de acordo com a alocação proporcional.

ANO	QUANTIDADE
1987	1
1989	3
1990	1
1991	2
1992	8
1993	42
1994	246
1995	213
1996	190
1997	204
1998	87
1999	129
2000	3
2001	1
2005	1
TOTAL	1131

#### 4. Identificação das amostras

Na última etapa do plano amostral, tem-se a coleta dos processos em cada estrato. Esta deve ser obtida de forma aleatória e sem reposição. Para isto utilizou-se um algoritmo computacional através do software R versão 4.1.3. No link abaixo é apresentada a lista de processos sorteados aleatoriamente e seus respectivos anos de distribuição:

[https://docs.google.com/spreadsheets/d/1o0rnnOLXZnXeJmQWVr4HdaoINlh\\_cRefBu1zZs1UxGY/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1o0rnnOLXZnXeJmQWVr4HdaoINlh_cRefBu1zZs1UxGY/edit?usp=sharing)

#### 5. Referências

- i. R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- ii. Conselho Nacional de Justiça (CNJ) Manual de Gestão Documental do Poder Judiciário / Conselho Nacional de Justiça. Programa Nacional de Gestão Documental e Memória do Poder Judiciário (Proname) – Brasília: CNJ, 2021.