

DETERMINAÇÃO DE DISTRIBUIÇÕES PROBABILÍSTICAS POR MEIO DO MÉTODO DE MONTE CARLO PARA O PROCESSO DE CHEGADA DE DOADORES EM UM BANCO DE SANGUE

IZABELLA BERNARDINELLI CESTARI

GILBERTO CLÓVIS ANTONELLI

Resumo

A doação de sangue é um processo complexo, pois envolve a questão humanitária e está mais atrelada ao emocional do que o racional. Devido ao seu caráter único, o mapeamento de processos de doação de sangue é difícil, pois vários fatores podem influenciar na decisão de um possível doador doar sangue. Este trabalho tem como objetivo encontrar o perfil de horário de chegada de possíveis doadores de sangue em um Banco de Sangue. As informações da quantidade de doadores que chegam de hora em hora extraídas do banco de dados serão estratificadas e servirão para determinar as distribuições probabilísticas pelo método de Monte Carlo. Com o auxílio da ferramenta ExpertFit® do Software FlexSim será possível definir quais equações descrevem o processo de chegada. Com o uso de ferramentas visuais, como gráficos, será possível analisar a existência ou não de padrões definidos na chegada de doadores associados as horas do dia e aos dias da semana. Com a conclusão destas análises será possível auxiliar o Banco de Sangue no gerenciamento de pessoal e realocação de funcionários visando a qualidade do atendimento e do serviço prestado.

Palavras-chave: *Pesquisa Operacional; Monte Carlo; Doação de Sangue; Distribuição de Probabilidade.*

1. Introdução

A doação de sangue é um processo complexo, pois envolve a questão humanitária e está mais atrelada ao emocional do que o racional. É devido a essa característica humanitária que mapear um processo de doação de sangue é difícil, pois os eventos não ocorrem de maneira constante e são diversos os fatores que podem influenciar a decisão de um possível doador comparecer a um banco de sangue e doar sangue.

Para Oswalt (1977), são vários os aspectos que conduzem uma pessoa a doar sangue: altruísmo/humanitarismo, programas de garantia, reposição, pressão social, necessidades da comunidade, recompensa e publicidade.

Ludwig e Rodrigues (2005) destacam a importância de hemocentros e serviços de hemoterapia conhecer as necessidades dos doadores para que assim possam aprimorar suas relações. Ao passo que as necessidades dos doadores se tornam mais arrojadas, aumentam as

exigências de respostas das instituições quanto à disponibilização de recursos humanos e materiais.

Programas de captação de doadores de sangue que objetivem a conscientização da doação voluntária devem atentar-se não apenas a quantidade de doadores, mas também para aprimorar o perfil das doações, garantindo elevado padrão de qualidade do sangue coletado e transfundido. (MINISTÉRIO DA SAÚDE, 2015).

Diante da necessidade de um grau elevado de qualidade nos serviços de coleta de sangue, se identificou a importância da realização de um estudo que caracterize um perfil de horário em que possíveis doadores chegam ao Banco de Sangue do Hospital do Câncer de Maringá.

O tema do estudo foi escolhido dado à falta de estudos aprofundados em temas relacionados ao serviço de hemoterapia. Devido ao seu caráter único, o serviço que o Banco de Sangue presta à sociedade deve ter seus processos estudados de forma mais próxima à realidade.

Com o perfil de horário que possíveis doadores chegam ao Banco de Sangue os gestores poderão tomar decisões que impactem diretamente na excelência do serviço prestado à comunidade.

O estudo em questão foi realizado no Banco de Sangue situado junto ao complexo do Hospital do Câncer de Maringá da cidade de Maringá, Paraná. O Banco de Sangue realiza cadastro, triagem clínica, triagem hematológica, coleta de sangue, análises sorológicas e imunohematológicas, e o armazenamento das bolsas que serão posteriormente utilizadas por hospitais da região de Maringá.

O estudo visa mostrar em quais horários do dia e em quais dias da semana há maior fluxo de chegada de possíveis doadores no Banco de Sangue. Através da análise do banco de dados da recepção do Banco de Sangue em um período pré-estabelecido e posterior interpretação dos dados por meio de gráficos e distribuições probabilísticas, espera-se uma correlação do perfil encontrado com outros fatores.

O objetivo geral deste trabalho fundamenta-se na estratificação de dados da chegada de possíveis doadores do Banco de Sangue através da determinação de distribuições probabilísticas pelo método de Monte Carlo, auxiliando em alterações no seu gerenciamento para que os serviços prestados tenham maior excelência.

Para que seja alcançado o objetivo principal do trabalho, deve ser necessário que durante o processo de desenvolvimento do mesmo sejam definidas ações como:

- Coletar informações do banco de dados utilizando o *Microsoft Power BI* para estratificar a ocorrência de chegada de possíveis doadores no Banco de Sangue.
- Determinar a distribuição probabilística dos dados do método de Monte Carlo usando a ferramenta *ExpertFit®* do *software FlexSim*;
- Analisar as distribuições determinadas pelo *software*.

A conclusão desses objetivos torna possível a determinação do perfil de horário de chegada dos possíveis doadores ao Banco de Sangue de forma mais próxima à realidade.

O desenvolvimento deste trabalho busca auxiliar no desenvolvimento de novos estudos sobre o tema, visto que não existem muitos estudos sobre os processos envolvidos em bancos de sangue ou hemocentros.

2. Fundamentação Teórica

2.1 Distribuição de probabilidade

Segundo Bateman *et al* (2013), uma distribuição de probabilidade é descrita como um conjunto de valores que relaciona a frequência que um evento ocorre ou virá a ocorrer. Qualquer processo que possua chegadas e saídas repetidamente, como chegadas de clientes em um sistema, pode ser descrito por uma distribuição de probabilidade. Para que se crie uma distribuição de probabilidade é preciso, segundo os autores, obter dados empíricos, ou seja, medidas coletadas no seu formato original. O próximo passo é criar um histograma de frequências relativas dos dados. Para isso, é necessário que se agrupe os dados em intervalos de classe. Os intervalos variam de cinco a vinte dependendo do número de dados da amostra. Quanto menor o número de dados, menor é o número de intervalos necessários. A frequência é dada pelo número de vezes que dados pertencentes a cada intervalo aparecem. A partir da divisão da frequência pelo número total de dados podemos obter a frequência relativa ou probabilidade. (BATEMAN *et al*, 2013)

Uma variável aleatória é uma variável cuja saída é determinada pelo resultado de um experimento, ou seja, muda de forma indefinida ao longo do tempo. A variável aleatória pode ser definida em contínua ou discreta, sendo a variável aleatória discreta aquela na qual a amplitude entre possíveis valores é finita e variável aleatória contínua aquela na qual a amplitude entre possíveis valores é infinita. Uma variável aleatória é geralmente representada

por um X maiúsculo e os valores associados a ela com um x minúsculo. (BATEMAN *et al*, 2013)

Bateman *et al* (2013) discorrem ainda sobre a função densidade de probabilidade, função na qual relaciona as probabilidades correspondentes aos valores x individuais associados com uma variável aleatória X. Encontrar a equação matemática que descreve uma função densidade, segundo os autores, é sempre difícil. Devido a essa dificuldade que frequentemente se tem uma aproximação da função encontrada com uma distribuição padrão conhecida (por exemplo: normal, exponencial, gama etc.) que representa com exatidão a distribuição de frequência relativa dos dados reais. A partir do formato do histograma de frequências relativas podemos obter dicas para encontrar uma distribuição padrão que represente os dados analisados.

Segundo Bussab e Morettin (2011) a distribuição padrão Normal é um modelo de distribuição probabilística fundamental usado em probabilidades e inferências estatísticas e que sua origem vem de trabalhos desenvolvidos por Gauss sobre erros de observações astronômicas. É possível então, dizer que uma variável contínua “X” possui distribuição Normal com parâmetros μ (média), σ (desvio-padrão) e σ^2 (variância) com a média indo de “menos” infinito até “mais” infinito ($-\infty < \mu < +\infty$) e a variância de 0 até infinito ($0 < \sigma^2 < \infty$) se a sua densidade for dada pela Equação 1.

$$f(x; \mu; \sigma^2) = \frac{1}{\sigma \sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, \quad -\infty < x < \infty \quad (1)$$

Denota-se a densidade da distribuição Normal da variável “X” simplesmente por $X \sim N(\mu, \sigma^2)$, em que N, representa a distribuição Normal.

A segmentação dos dados em subconjuntos é dada pelos intervalos de classes, ou seja, seguindo um critério estabelecido se separa os dados em conjuntos com o propósito de simplificar a observação dos mesmos.

Sturges (2012) apresentou um método para calcular a amplitude de um conjunto de dados. Segundo o autor, no caso de uma única série estatística de dados de variáveis aleatórias de alcance R - representa a amplitude do conjunto de dados - com N itens envolvidos em uma simulação computacional, a amplitude da classe de intervalos ótima para a determinação de distribuição de probabilidades pode ser dada pela Equação 2:

$$C = \frac{R}{1 + 3,332 \log N} \quad (2)$$

A Equação 2 apresenta o intervalo de classe para o cálculo das médias, mediana, frequência, assimetrias, distribuições de frequência, entre outros. Utilizando as Equações 1 e 2 é possível realizar a determinação de distribuição de probabilidade pelo método analítico.

2.2 Método de Monte Carlo

O método de Monte Carlo é definido como uma ferramenta que gera dados artificiais a partir de um gerador de números aleatórios. A partir dos dados artificiais se origina distribuições de probabilidades para cada um dos dados. São essas distribuições que tornam possíveis a simulação de processos estocásticos, ou seja, que acontecem ao acaso. (MENOSSE, 2002)

Segundo Bindel e Goodman (2009), o método de Monte Carlo é utilizado por causa do problema de dimensionamento. O trabalho para resolver um problema em muitas dimensões pode crescer de forma exponencial com a dimensão. É por isso que muitos problemas de alta dimensão podem ser resolvidos, até onde se sabe, apenas pelo método de Monte Carlo.

2.3 Software ExpertFit®

Baseado no método de Monte Carlo, o *software ExpertFit®* tem a função de encontrar, por meio de conjuntos de dados inseridos no sistema, modelos de distribuição probabilísticos que melhor descrevem os mesmos.

Através de um algoritmo de busca, o *software*, desenvolvido por *Averill M. Law & Associates, Inc*, nos Estados Unidos, tem por objetivo encontrar de forma automática o ajuste do conjunto de dados à melhor distribuição de probabilidade. (EXPERTFIT, 2019).

O *software ExpertFit®* pode ser encontrado associado em alguns *softwares* de simulação, como por exemplo o *FlexSim*, e permite a criação de histogramas baseado nas curvas de probabilidades encontradas, além de gerar um gráfico comparando as distribuições. O *software* assegura a inserção de até 100.000 dados, com 40 distribuições de probabilidade disponíveis para análise. Na ausência de dados é possível a geração de valores aleatórios a partir de qualquer distribuição de probabilidade. (EXPERTFIT, 2019).

2.4 Comportamento de Filas

A formação de filas é algo que acontece com frequência no cotidiano das pessoas, seja no supermercado, lotéricas, bancos ou até mesmo nas praças de pedágio. Quando a procura por um serviço é maior que a capacidade do processo de atender, a fila é formada. A teoria de

filas busca, sempre que possível, auxiliar e aprimorar o sistema de filas objetivando um equilíbrio no sistema a fim de que os custos e o tempo de espera do cliente sejam menores. (LAMARCA, 2016).

Segundo Arenales *et al* (2007) existem vários tipos de filas, dentre eles:

- Fila única e um servidor;
- Fila única e múltiplos servidores em paralelo;
- Múltiplas filas e múltiplos servidores em paralelo;
- Fila única e múltiplos servidores em série.

Os três primeiros tipos de filas citados dispõem apenas uma etapa a ser realizada, por exemplo: posto de combustível, caixas de banco, entre outros. O modelo “fila única e múltiplos servidores em série” é diferente, pois, apresenta várias etapas a serem realizadas em série, ou seja, uma etapa após a outra. O quarto modelo de fila é comum em lugares como: hospitais, pronto socorro, *drive-in* de *fast-food*, entre outros. (ARENALES *et al*, 2007).

O processo de fila pode ser estocástico ou determinístico. Os processos que se desenvolvem de acordo com as leis de probabilidade e podem ser representados por variáveis aleatórias são os chamados estocásticos. São chamados de determinísticos os processos onde é possível determinar exatamente o tempo que o mesmo irá consumir. Além disso, há outras questões pertinentes: se os clientes chegam de forma individual ou se chegam vários clientes ao mesmo tempo, se tem a possibilidade do cliente desistir da fila, se tem a possibilidade do cliente mudar de uma fila para outra, se a eficiência do serviço prestado varia conforme o tamanho da fila. (LAMARCA, 2016).

2.5 Distribuição de Chegada

Um modelo de filas pode ser descrito pela notação Kendall-Lee: **A/B/c/K/m/Z**.

As notações “A” e “B” indicam qual é o tipo de distribuição que os intervalos entre chegadas e tempo de serviço, respectivamente, seguem (Marciviana, Erlang, hiper-exponencial, determinística ou geral). A disciplina da fila é indicada por “Z”. “K” representa a capacidade máxima do sistema, ou seja, número máximo de clientes permitidos no sistema. O “m” se refere ao tamanho da população que pode fornecer clientes ao sistema, finito ou infinito. O “C” indica o número de servidores que estão inseridos no sistema. É usual que se

omita “K” e “m” da notação caso sejam infinitos. Caso a disciplina seja de uma fila genérica, o “Z” também é omitido. (PRADO, 2009).

3. Metodologia

De acordo com Prodanov e Freitas (2013), a pesquisa pode ser classificada quanto à sua natureza, objetivos, procedimentos e forma de abordagem do problema.

Quanto à natureza, este trabalho pode ser classificado como pesquisa aplicada, pois tem o objetivo de gerar conhecimentos de aplicação prática para a solução de problemas, envolvendo interesses reais. Do ponto de vista dos objetivos, a pesquisa é descritiva, irá descrever as características de determinada população ou fenômeno. A pesquisa realizada neste trabalho é vista como um estudo de caso, pois envolve o estudo profundo e exaustivo de um ou poucos objetos visando o seu amplo e detalhado conhecimento. A abordagem do problema é classificada como quantitativa, pois se busca traduzir em números opiniões e informações para classificá-las e analisá-las, utilizando recursos e técnicas estatísticas. (PRODANOV e FREITAS, 2013).

Como amostra de estudo, serão utilizadas as pessoas e processos envolvidos e existentes para doação de sangue do Banco de Sangue do Hospital do Câncer de Maringá. Assim os passos identificados para a realização deste estudo são:

- Revisão bibliográfica dos conceitos relacionados ao tema;
- Obter o banco de dados do Banco de Sangue junto ao mesmo;
- Extrair as informações do banco de dados;
- Estratificar as informações extraídas;
- Calcular, por meio do *software ExpertFit®*, as medidas necessárias para gerar a distribuição de probabilidade através dos dados estratificados;
- Ajustar os dados no *software ExpertFit®*;
- Gerar um gráfico de comparação entre as melhores funções de probabilidade no *software ExpertFit®*;
- Analisar as distribuições geradas;
- Conclusão do estudo com recomendações e ajustes para estudos futuros.

4. Resultados

Através da análise dos dados de atendimento do Banco de Sangue adquiridos na plataforma *Microsoft Power BI*, foi possível auferir a quantidade de pessoas que doam sangue diariamente. Com a estratificação dos dados obteve-se as quantidades de doares em períodos e dias pré-estabelecidos com o intuito de encontrar o perfil de horário.

Na Tabela 1 podemos analisar a quantidade de doares nos períodos manhã, tarde e noite nos meses de janeiro a outubro.

Tabela 1: Quantidade de doadores em três intervalos de tempo diferentes.

Mês	Período		
	01/01/2019	31/10/2019	
	Horário		
	07:00 - 10:59	11:00 - 12:59	13:00 - 18:00
Janeiro	184	126	300
Fevereiro	119	58	216
Março	143	68	175
Abril	191	87	240
Maio	138	75	232
Junho	119	49	159
Julho	133	64	256
Agosto	137	76	227
Setembro	141	59	235
Outubro	205	92	256
Total	1510	754	2296

Fonte: A Autora (2019).

O intuito dessa medição é analisar qual a quantidade de doares que comparecem no Banco de Sangue na parte da manhã, na hora do almoço e a na parte da tarde e se as distribuições desses três períodos seguem algum padrão.

Utilizando os dados da Tabela 1 pode-se realizar o cálculo para a construção da Tabela 2, onde se tem a quantidade média de doadores por hora dada em cada intervalo de tempo no período de janeiro a outubro. O cálculo foi empreendido com a finalidade de comparar as distribuições dos três intervalos.

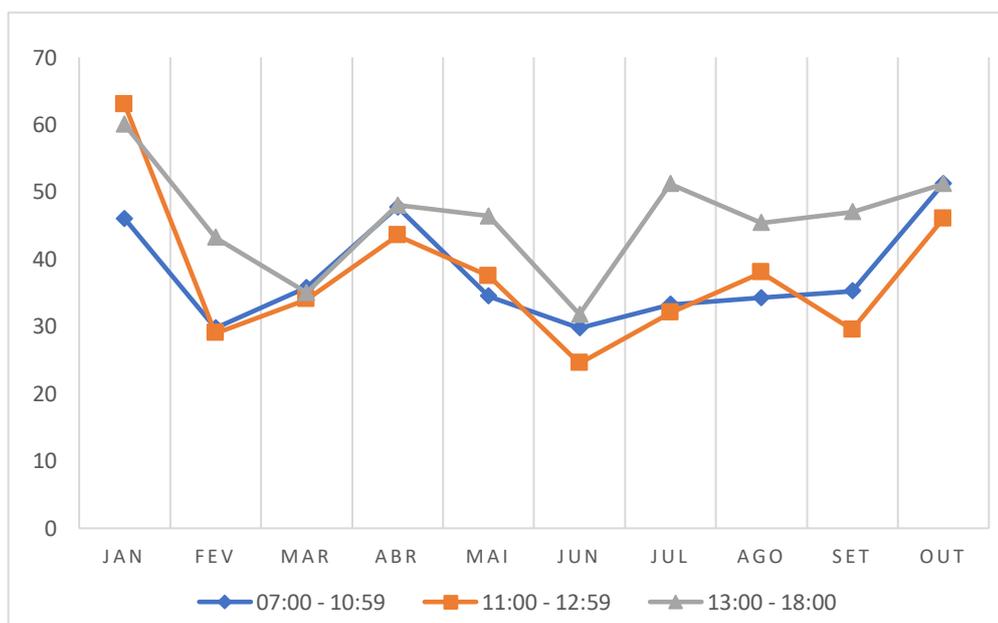
Tabela 2: Quantidade de doadores por hora em três intervalos de tempo diferentes.

Mês	Período 01/01/2019 31/10/2019		
	Horário		
	07:00 - 10:59	11:00 - 12:59	13:00 - 18:00
Janeiro	46	63	60
Fevereiro	29,75	29	43,2
Março	35,75	34	35
Abril	47,75	43,5	48
Maio	34,5	37,5	46,4
Junho	29,75	24,5	31,8
Julho	33,25	32	51,2
Agosto	34,25	38	45,4
Setembro	35,25	29,5	47
Outubro	51,25	46	51,2
Total	377,5	377	459,2

Fonte: A Autora (2019).

Para a construção da Figura 1 foram utilizados os dados da Tabela 2.

Figura 1: Quantidade de doadores por hora em três intervalos de tempo diferentes.

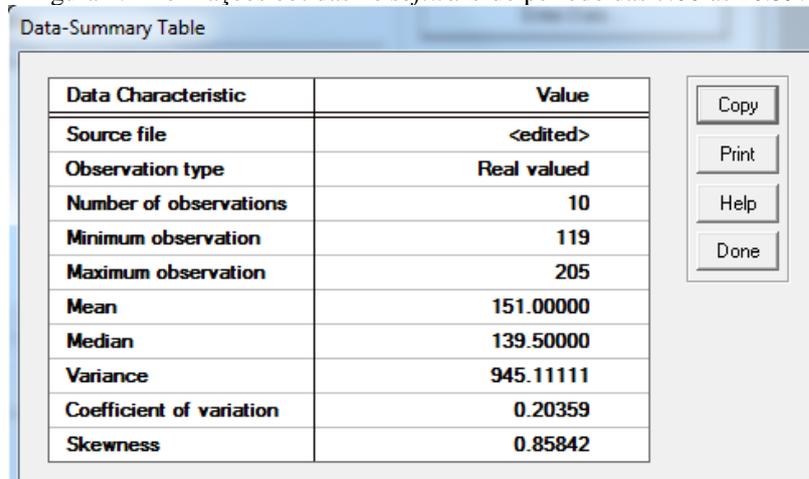


Fonte: A Autora (2019).

Ao analisar a Figura 1 pode-se perceber um pequeno padrão no comportamento durante os meses do ano. Ao realizar a análise dos períodos percebe-se que o período da tarde é o que predominantemente recebe mais doadores. Já no período da manhã e na hora do almoço a quantidade média de doadores que chegam ao Banco de Sangue se alternam durante os meses do ano. Ao analisar os dados anualmente percebe-se que há picos de quantidade de doadores nos meses de férias escolares.

Ao incluir os dados da primeira coluna da Tabela 1 na ferramenta *ExpertFit*® obtemos os resultados da Figura 2.

Figura 2: Informações obtidas no *software* do período das 7:00 às 10:59.



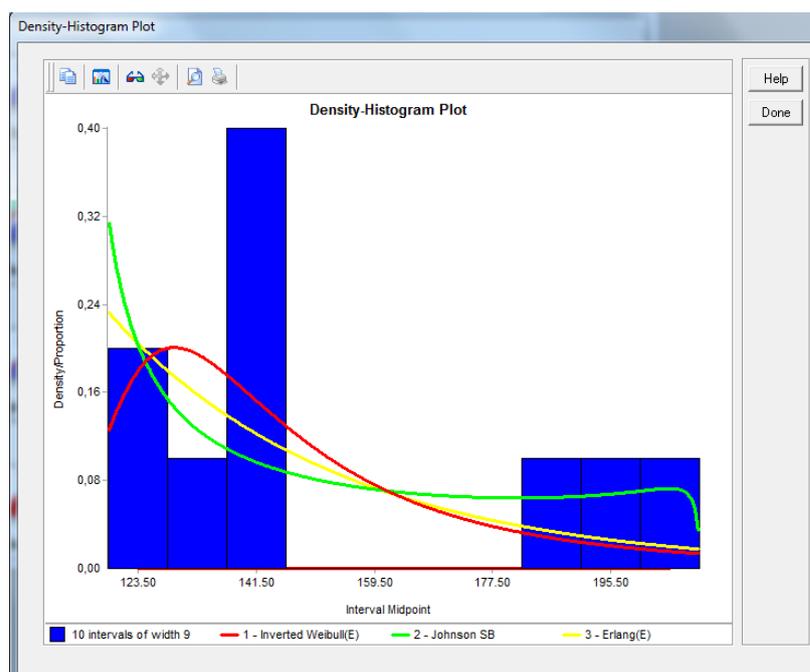
Data Characteristic	Value
Source file	<edited>
Observation type	Real valued
Number of observations	10
Minimum observation	119
Maximum observation	205
Mean	151.00000
Median	139.50000
Variance	945.11111
Coefficient of variation	0.20359
Skewness	0.85842

Fonte: A Autora (2019).

Têm-se: número de observações (10), menor valor observado (119), maior valor observado (205), média (151), mediana (139,50), variância (945,11), coeficiente de variância (0,20359) e assimetria (0,85842). Com as medidas geradas para o período da manhã (7:00 às 10:59) se torna viável a determinação da distribuição de probabilidade para o período.

Ao realizar o ajuste dos dados da Figura 2 na ferramenta *ExpertFit*® obtém-se as três distribuições de probabilidade com o melhor Score Relativo, ou seja, melhor nota, que se pode analisar na Figura 3.

Figura 3: Histograma do período das 7:00 às 10:59.

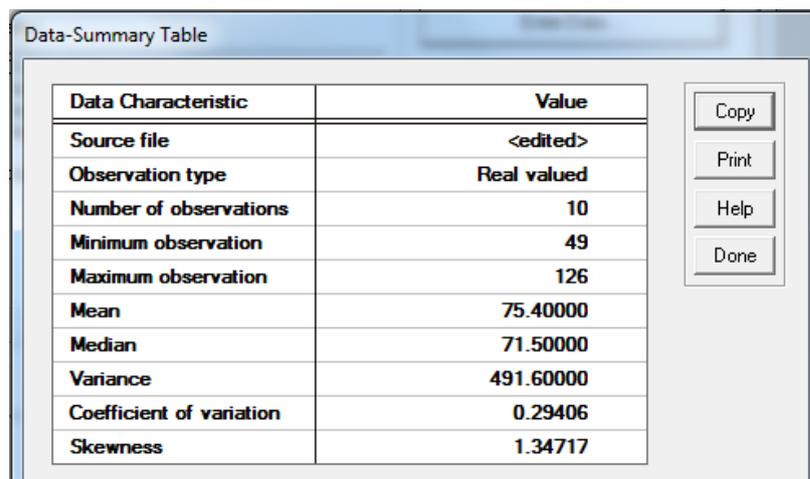


Fonte: A Autora (2019).

Para o período das 7:00 às 10:59 as três melhores distribuições foram as funções Weibull, Johnson SB e Erlang. Na Figura 3 tem-se o gráfico comparativo das três funções em que se observa algumas diferenças entre as curvas.

A inserção dos dados da segunda coluna da Tabela 1 na ferramenta *ExpertFit*® gera os resultados da Figura 4 a seguir.

Figura 4: Informações obtidas no software do período das 11:00 às 12:59.



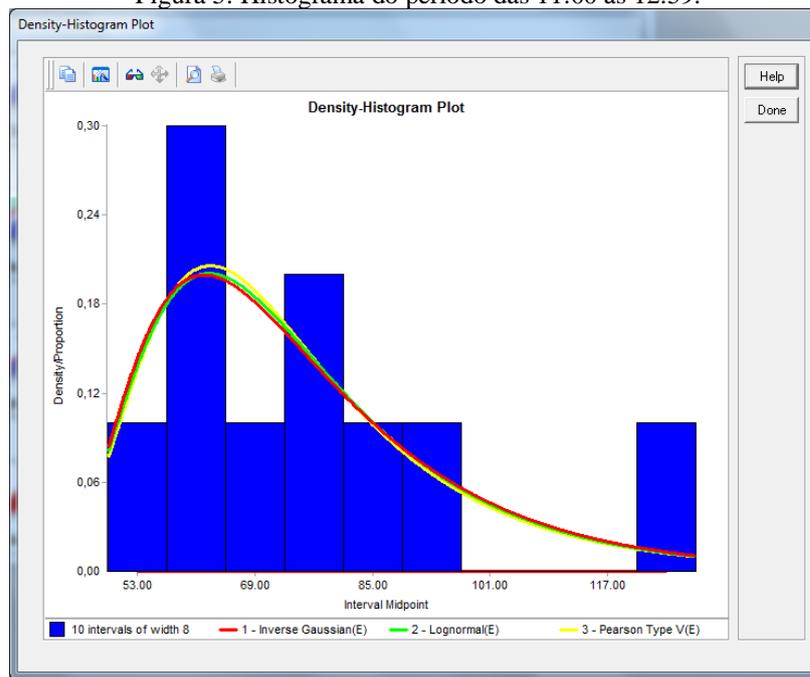
Data Characteristic	Value
Source file	<edited>
Observation type	Real valued
Number of observations	10
Minimum observation	49
Maximum observation	126
Mean	75.40000
Median	71.50000
Variance	491.60000
Coefficient of variation	0.29406
Skewness	1.34717

Fonte: A Autora (2019).

Utilizando-se os resultados da Figura 4 torna-se executável a determinação da distribuição de probabilidade para o período das 11:00 às 12:59 (hora do almoço).

Com o ajuste na ferramenta *ExpertFit*® se chega as três distribuições de probabilidade com a melhor nota, dada pelo *software* como Score Relativo. Na Figura 5 têm-se o gráfico onde é comparado as três melhores distribuições para o período das 11:00 às 12:59.

Figura 5: Histograma do período das 11:00 às 12:59.



Fonte: A Autora (2019).

As três melhores distribuições para a hora do almoço são: Gaussiana (Normal), Log-Normal e Pearson. Com a representação pode-se observar a semelhança entre as curvas das distribuições determinadas.

Inserindo os dados da terceira coluna da Tabela 1 na ferramenta *ExpertFit*® é gerada a Figura 6.

Figura 6: Informações obtidas no software do período das 13:00 às 18:59.

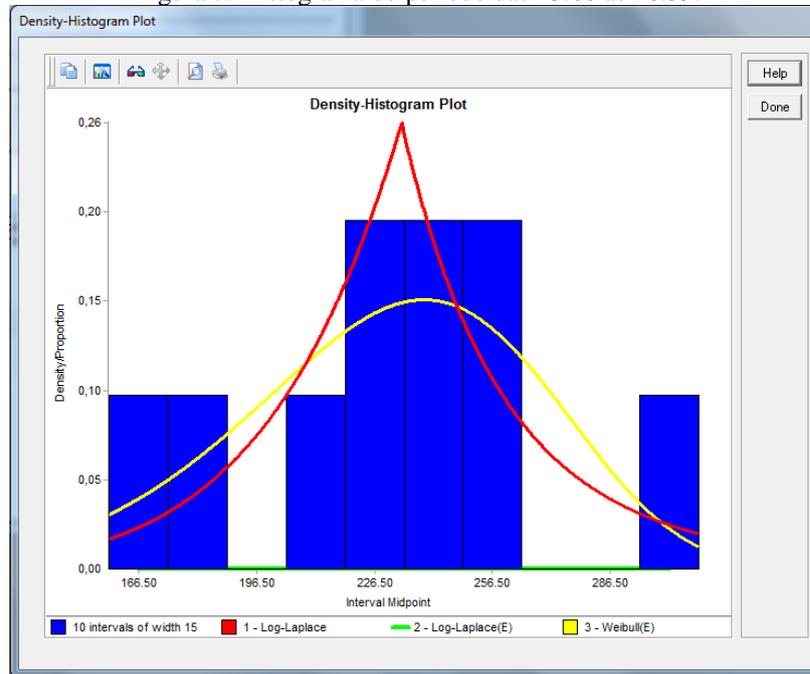
Data Characteristic	Value
Source file	<edited>
Observation type	Real valued
Number of observations	10
Minimum observation	159
Maximum observation	300
Mean	229.60000
Median	233.50000
Variance	1.627.82222
Coefficient of variation	0.17572
Skewness	-0.27532

Fonte: A Autora (2019).

Com os resultados da Figura 6 é possível realizar o ajuste dos dados na ferramenta *ExpertFit*® e obter as três distribuições de probabilidade com o melhor Score Relativo, ou

seja, melhor nota. As distribuições de probabilidades obtidas para o período das 13:00 às 18:59 pode ser analisada na Figura 7.

Figura 7: Histograma do período das 13:00 às 18:59.



Fonte: A Autora (2019).

Observando a Figura 7 pode-se perceber que as três curvas das distribuições probabilísticas (*Laplace, Laplace e Weibull*) determinadas são bem diferentes entre si.

Na tabela 3 pode-se visualizar a melhor equação para cada intervalo de tempo e seu respectivo Score Relativo.

Tabela 3: Equação matemática para os intervalos de tempo manhã, almoço e tarde.

Intervalo	Nome Função	Parâmetros	Equação Matemática	Score Relativo	
7:00 - 11:00	Iverted Weibull	Location	85.837494	ivertedweibull (85.837494, 48.227468, 2.762528, <stream>)	85,48
		Scale	48.227468		
		Shape	2.762528		
11:00 - 13:00	Normal (Gaussiana)	Location	33.657284	iversegaussian (33.657284, 41.742716, 159.362886, <stream>)	89,06
		Scale	41.742716		
		Shape	159.362886		
13:00 - 18:00	Laplace	Location	0.000000	loglaplace (0.000000, 233.495182, 7.976353, <stream>)	99,19
		Scale	233.495182		
		Shape	7.976353		

Fonte: A Autora (2019).

Os parâmetros são os mesmos para as três funções: Location (localização), Sacale (escala) e Shape (forma). O parâmetro Location fornece o posicionamento da distribuição,

Scale define a amplitude da distribuição e Shape define qual será a forma da distribuição, ou seja, qual função. O parâmetro Stream se refere a uma variável aleatória que servirá de semente para que seja possível o cálculo da função.

Na Tabela 4 foi estratificado os dados das quantidades de doadores que chegam de hora em hora durante a semana (segunda à sexta-feira) no Banco de Sangue.

Tabela 4: Quantidade de doadores de hora em hora de segunda à sexta-feira.

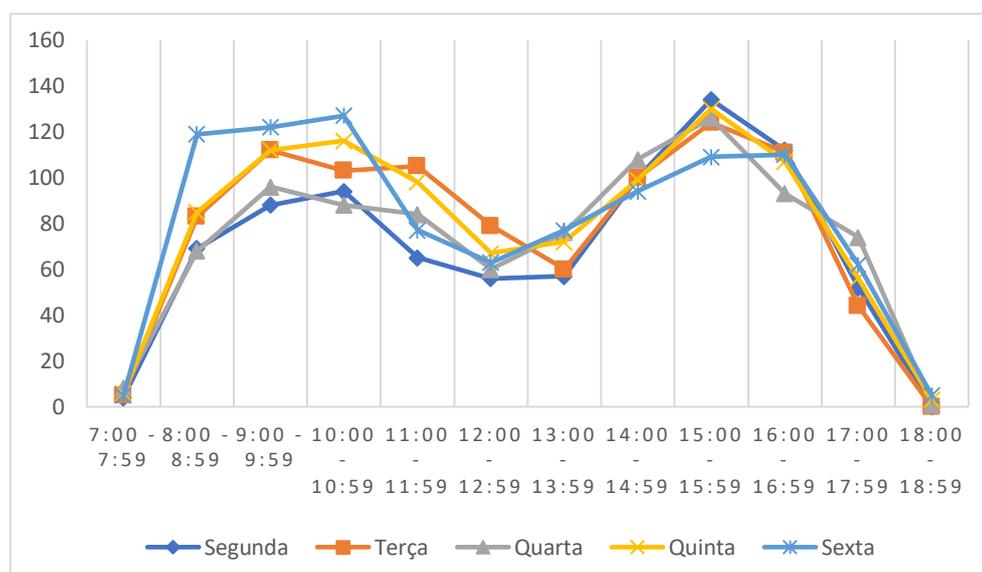
Horario	Período				
	01/01/2019	01/01/2019	01/01/2019	01/01/2019	01/01/2019
	Segunda	Terça	Quarta	Quinta	Sexta
7:00 - 7:59	4	5	8	6	5
8:00 - 8:59	69	83	68	85	119
9:00 - 9:59	88	112	96	112	122
10:00 - 10:59	94	103	88	116	127
11:00 - 11:59	65	105	84	98	77
12:00 - 12:59	56	79	60	67	63
13:00 - 13:59	57	60	76	72	77
14:00 - 14:59	100	100	108	99	94
15:00 - 15:59	134	124	126	130	109
16:00 - 16:59	112	111	93	107	110
17:00 - 17:59	52	44	74	56	62
18:00 - 18:59	0	0	1	3	5
TOTAL	831	926	882	951	970

Fonte: A Autora (2019).

Com esses dados torna-se possível a construção de um gráfico onde pode-se analisar se existe algum padrão relacionado à chegada de possíveis doadores no começo da manhã (quando o Banco de Sangue abre) na hora do almoço e no final da tarde, que é quando o Banco de sangue encerra as atividades.

Na Figura 8 a seguir pode-se notar um padrão.

Figura 8: Quantidade de doadores de hora em hora de segunda à sexta-feira.



Fonte: A Autora (2019).

A quantidade de doares é ampliada conforme o expediente começa e diminui conforme a hora do almoço se aproxima. Na parte da tarde ocorre um padrão parecido. A quantidade de doares cresce no início da tarde, tem um pico bem definido às 15 horas e é reduzida próxima às 18 horas.

Na Tabela 5 se encontra os dados de quantidade de doares que comparecem ao Banco de Sangue durante os dias da semana nos meses de janeiro a outubro.

Tabela 5: Quantidade de doadores nos meses de janeiro a outubro de segunda à sexta-feira.

Período	01/01/2019 - 31/10/2019					TOTAL
	Dia da semana					
Mês	Segunda	Terça	Quarta	Quinta	Sexta	TOTAL
Janeiro	105	162	129	130	84	610
Fevereiro	74	86	81	68	84	393
Março	71	59	70	64	122	386
Abril	116	111	62	147	82	518
Maio	58	72	82	115	118	445
Junho	86	61	55	39	86	327
Julho	91	74	128	89	71	453
Agosto	61	92	79	98	110	440
Setembro	75	73	97	102	88	435
Outubro	94	136	99	99	125	553
Total	831	926	882	951	970	4560

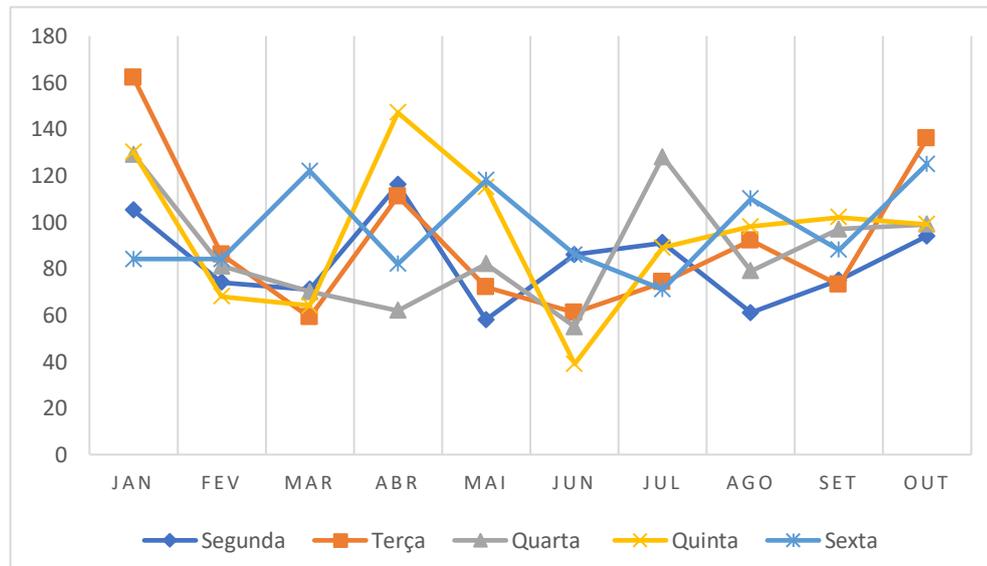
Fonte: A Autora (2019).

Ao plotar os dados em um gráfico foi possível verificar se há uma correlação entre a chegada de possíveis doadores de sangue no Banco de Sangue e o dia da semana. Por

exemplo, por segunda-feira e sexta-feira estarem próximas ao final de semana, é de praxe que as pessoas pensem que o fluxo de chegada de pessoas nesses dois dias da semana é maior, pois, no sábado o expediente é reduzido pela metade e no domingo não há expediente.

No Figura 9 foi plotado os dados da quantidade de doares que comparecem ao Banco de Sangue de segunda à sexta mensalmente.

Figura 9: quantidade de doares durante os dias da semana de janeiro a outubro.

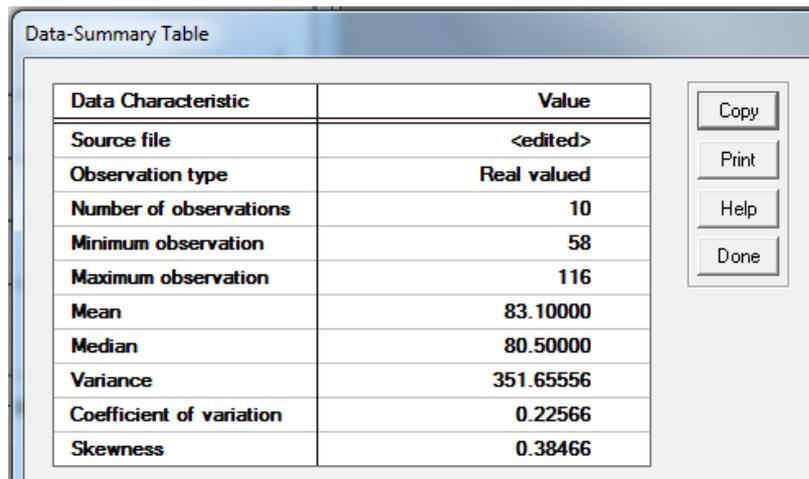


Fonte: A Autora (2019).

Na Figura 9 nota-se a falta de padrão. É possível observar que cada dia da semana teve seu pico em um dado mês. Como ocorreu com as terças-feiras em janeiro, quintas-feiras em abril e quartas-feiras em julho. Com a análise desse gráfico percebe-se que a quantidade de doares que chegam ao Banco de Sangue não está relacionada com os dias da semana. Dessa forma, pode-se concluir que a chegada de possíveis doadores independe do dia da semana.

Com a inserção os dados da primeira coluna da Tabela 5 na ferramenta *ExpertFit®* foram geradas as medidas na Figura 10 a seguir.

Figura 10: Informações obtidas no software para segunda-feira.



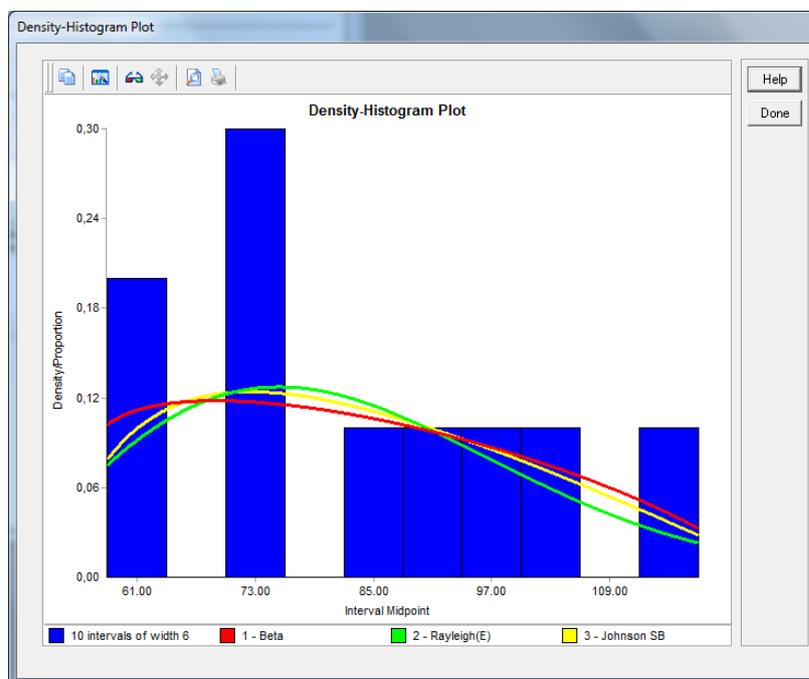
Data Characteristic	Value
Source file	<edited>
Observation type	Real valued
Number of observations	10
Minimum observation	58
Maximum observation	116
Mean	83.10000
Median	80.50000
Variance	351.65556
Coefficient of variation	0.22566
Skewness	0.38466

Fonte: A Autora (2019).

Usando os resultados da Figura 10 pode-se determinar a distribuição de probabilidade para as segundas-feiras de janeiro a outubro.

Ajustando as medidas no *software ExpertFit®* consegue-se obter as três distribuições de probabilidade com a melhor nota, ou seja, melhor Score Relativo que se pode analisar na Figura 11 abaixo.

Figura 11: Histograma dos dados mensais de segunda-feira.

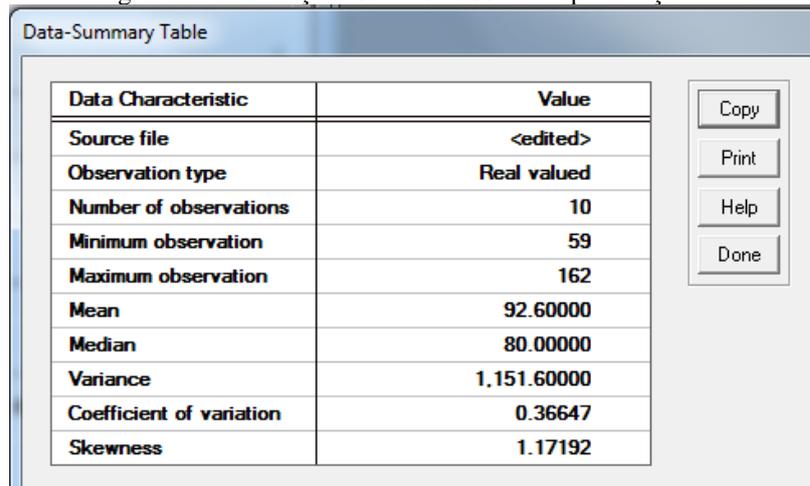


Fonte: A Autora (2019).

Para as segundas-feiras foi determinada as seguintes distribuições de probabilidades: Beta, Rayleigh e Johnson SB. As curvas de cada função podem ser vistas na Figura 11 acima. Com o gráfico de comparação fica nítido a semelhança que existe entre as três curvas.

As distribuições de probabilidade para as terças-feiras de janeiro a outubro serão determinadas através da inserção dos dados da Figura 12 no *software*.

Figura 12: Informações obtidas no software para terça-feira.

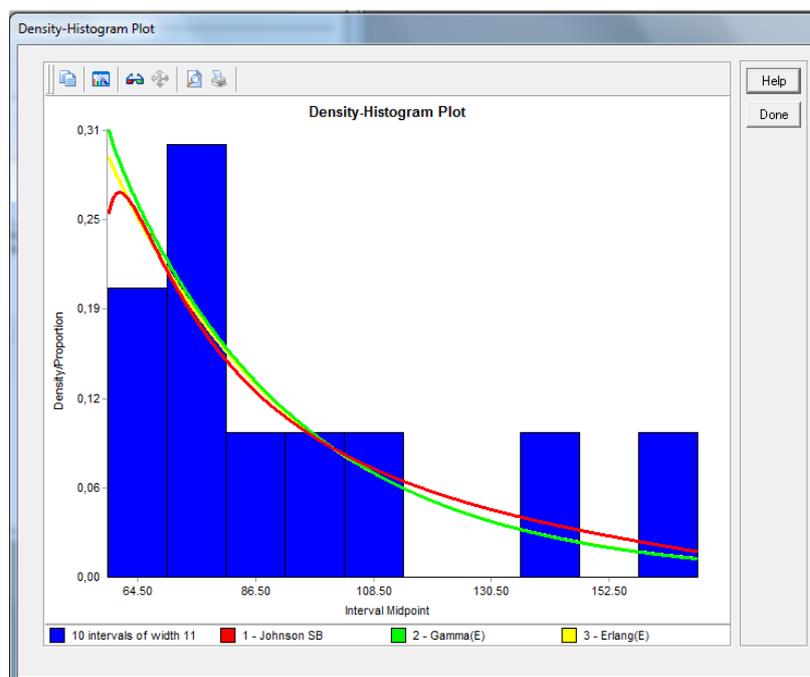


Data Characteristic	Value
Source file	<edited>
Observation type	Real valued
Number of observations	10
Minimum observation	59
Maximum observation	162
Mean	92.60000
Median	80.00000
Variance	1.151.60000
Coefficient of variation	0.36647
Skewness	1.17192

Fonte: A Autora (2019).

Ao ajustar os dados no *software ExpertFit®* é possível obter as três melhores distribuições de probabilidade (máximo Score Relativo).

Figura 13: Histograma dos dados mensais de terça-feira.

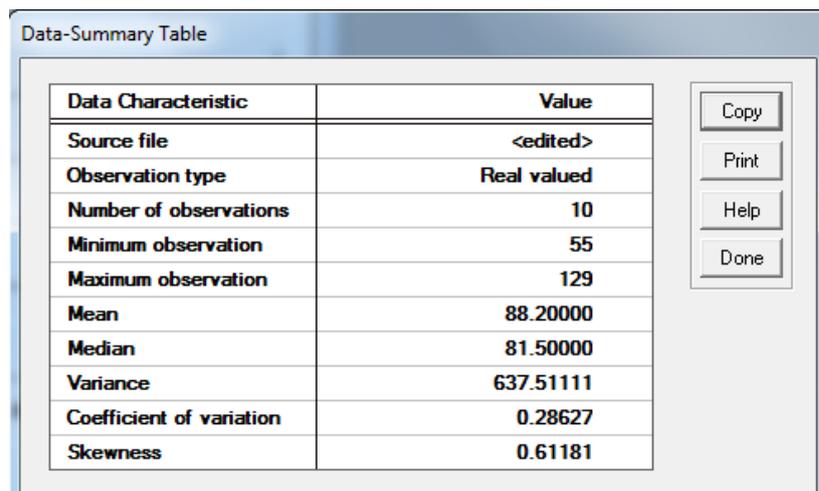


Fonte: A Autora (2019).

As distribuições de probabilidades determinadas para as terças-feiras foram: Johnson SB, Gamma e Erlang. Através do gráfico de comparação consegue-se visualizar a semelhança que existe entre as curvas.

Através da inserção dos dados da terceira coluna da Tabela 5 na ferramenta *ExpertFit*® foram gerados os resultados da Figura 14.

Figura 14: Informações obtidas no software para quarta-feira.



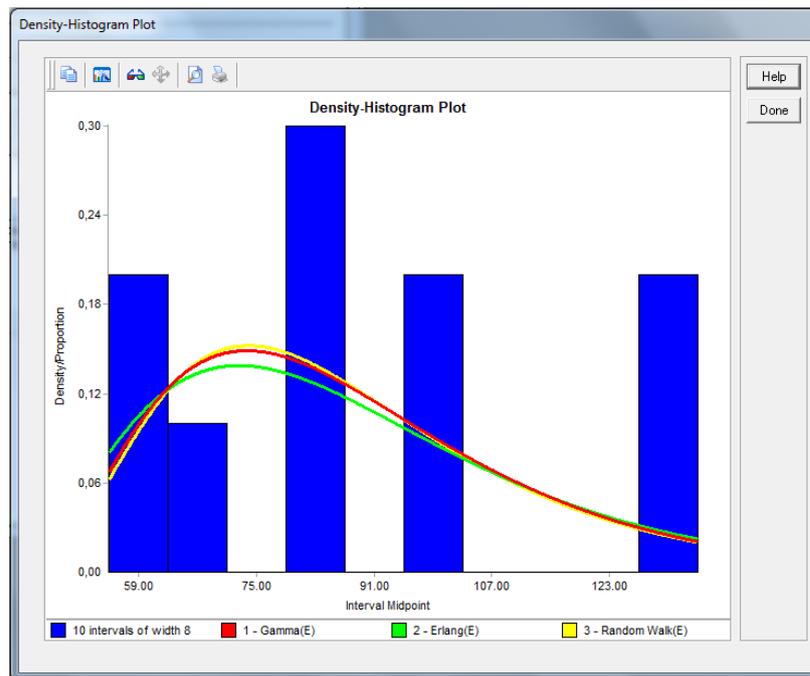
Data Characteristic	Value
Source file	<edited>
Observation type	Real valued
Number of observations	10
Minimum observation	55
Maximum observation	129
Mean	88.20000
Median	81.50000
Variance	637.51111
Coefficient of variation	0.28627
Skewness	0.61181

Fonte: A Autora (2019).

Utilizando os resultados da Figura 14 será possível a determinação das distribuições de probabilidade para as quartas-feiras de janeiro a outubro.

Com o auxílio da função ajuste da ferramenta *ExpertFit*® consegue-se atingir as melhores distribuições de probabilidade ranqueadas pelo Score Relativo, ou seja, pela melhor nota.

Figura 15: Histograma dos dados mensais de quarta-feira.



Fonte: A Autora (2019).

As três melhores distribuições de probabilidades determinadas pelo software para as quartas-feiras foram: Gamma, Erlang e Random Walk (caminhada aleatória). Ao gerar o gráfico de comparação pode-se ver três curvas idênticas.

Com os dados da quinta coluna da Tabela 5 inseridos na ferramenta *ExpertFit*® obtivemos as medidas da Figura 16 a seguir.

Figura 16: Informações obtidas no software para quinta-feira.

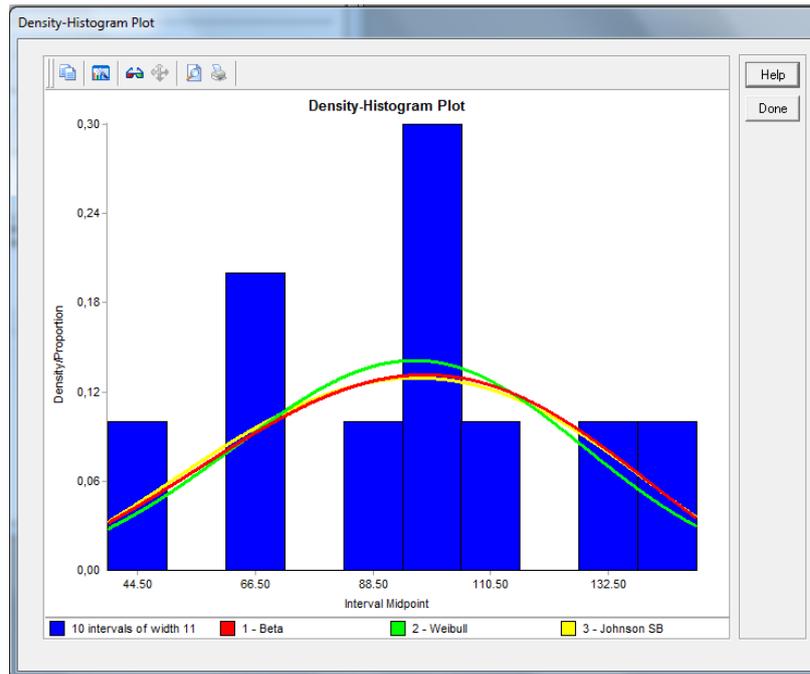
Data Characteristic	Value
Source file	<edited>
Observation type	Real valued
Number of observations	10
Minimum observation	39
Maximum observation	147
Mean	95.10000
Median	98.50000
Variance	1,029.43333
Coefficient of variation	0.33738
Skewness	-0.15209

Fonte: A Autora (2019).

Com o uso dos dados (Figura 16) encontrados no *software*, a determinação das distribuições probabilísticas para as quintas-feiras (de janeiro a outubro) se tornam uma tarefa simples.

Após executar a função de ajuste de dados do *software ExpertFit®* nos resultados, têm-se a Figura 17.

Figura 17: Histograma dos dados mensais de quinta-feira.

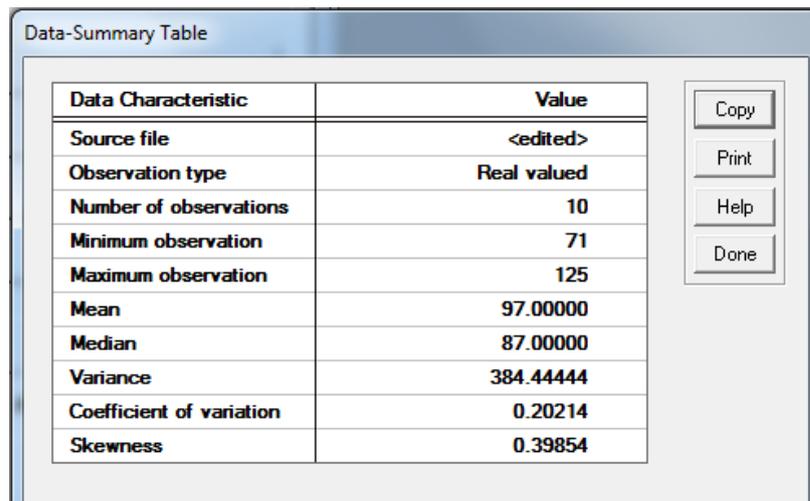


Fonte: A Autora (2019).

O gráfico de comparação apresentou as três distribuições de probabilidade que tiveram maior Score Relativo, sendo: Beta, Weibull e Johnson SB. Com o mesmo podemos perceber o quão semelhantes essas curvas são.

Utilizando a função dados da ferramenta *ExpertFit®*, foram observados os resultados da Figura 18 para os dados da sexta coluna da Tabela 5.

Figura 18: Informações obtidas no software para sexta-feira.



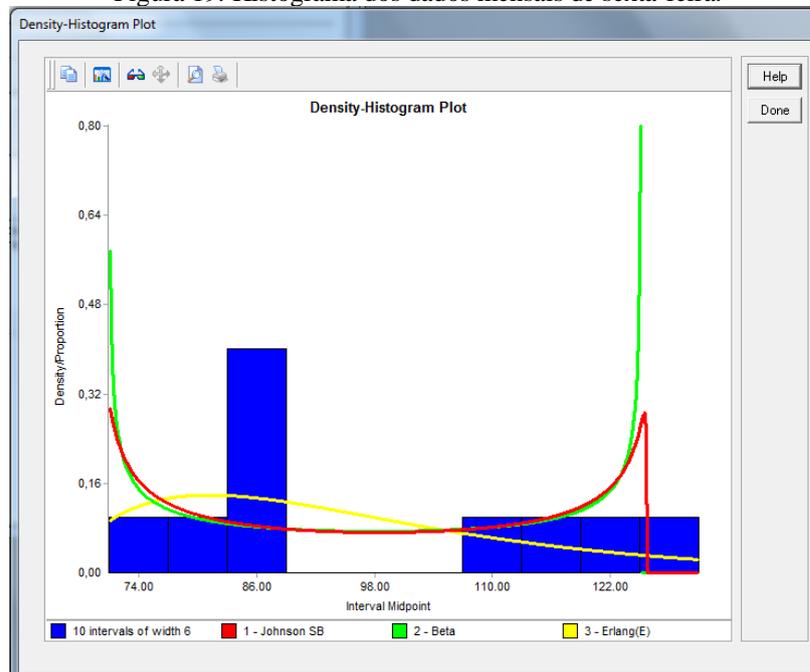
Data Characteristic	Value
Source file	<edited>
Observation type	Real valued
Number of observations	10
Minimum observation	71
Maximum observation	125
Mean	97.00000
Median	87.00000
Variance	384.44444
Coefficient of variation	0.20214
Skewness	0.39854

Fonte: A Autora (2019).

Por meio dos resultados da Figura 18, a determinação da distribuição de probabilidade para as sextas-feiras torna-se descomplicada.

Empregando os dados na função ajuste da ferramenta *ExpertFit*®, o mesmo apontou três distribuições de probabilidade como as melhores.

Figura 19: Histograma dos dados mensais de sexta-feira.



Fonte: A Autora (2019).

As três melhores distribuições foram as seguintes: Johnson SB, Beta e Erlang. As mesmas obtiveram o maior Score Relativo entre as funções disponíveis na ferramenta.

Mediante Figura 19 acima, pode-se fazer a comparação das curvas das funções apontadas. O gráfico deixa visível que as curvas têm diferenças entre si.

Da mesma forma que foi feito para as funções dos três intervalos de tempo, temos a Tabela 6 com a melhor distribuição para cada dia da semana gerada no *ExpertFit*® com seu respectivo Score Relativo, ou seja, sua porcentagem de correlação.

Tabela 6: Equação matemática para os intervalos de tempo manhã, almoço e tarde.

Dia da Semana	Nome Função	Parâmetros		Equação Matemática	Score Relativo
Segunda	Beta	Minimum	54.699311	beta (54.699311, 124.607086, 1.186054, 1.723907, <stream>)	97,66
		Maximum	124.607086		
		Shape1	1.186054		
		Shape2	1.723907		
Terça	Johnson SB	Minimum	55.656331	johnsonbounded (55.656331, 201.504499, 1.063880, 0.720290, <stream>)	89,52
		Maximum	201.504499		
		Shape1	1.063880		
		Shape2	0.720290		
Quarta	Gamma	Location	44.631890	gamma (44.631890, 14.475087, 3.009869, <stream>)	90,32
		Scale	14.475087		
		Shape	3.009869		
Quinta	Beta	Minimum	1.341553	beta (1.341553, 172.679092, 3.831505, 3.171379, <stream>)	97,5
		Maximum	172.679092		
		Shape1	3.831505		
		Shape2	3.171379		
Sexta	Johnson SB	Minimum	70.489947	johnsonbounded (70.489947, 125.835403, 0.021000, 0.422688, <stream>)	89,52
		Maximum	125.835403		
		Shape1	0.021000		
		Shape2	0.422688		

Fonte: A Autora (2019).

Os parâmetros para as funções de segunda, terça, quinta e sexta são os mesmos: Minimum (mínimo), Maximum (máximo), Shape1 (forma1) e Shape2 (forma2). Para a função de quarta-feira os parâmetros são Location (localização), Scale (escala) e Shape (forma).

Os parâmetros Minimum e Maximum geram, respectivamente, os valores mínimos e máximos que função pode atingir. Seu cálculo é feito por meio da adição e subtração do desvio padrão na média dos dados. Os parâmetros Shape1 e Shape2 referem-se à forma que a distribuição irá tomar, ou seja, qual função.

O parâmetro Location concede a posição da distribuição, Scale estabelece a amplitude da distribuição e Shape indica qual será a forma da distribuição. O parâmetro Stream é dado

como uma variável semente, ou seja, é uma variável aleatória que servirá para dar o início no cálculo da função.

Na Tabela 7 tem-se os dados da quantidade de doares que comparecem ao Banco de Sangue de hora em hora aos Sábados.

Tabela 7: Quantidade de doares de hora em hora no sábado.

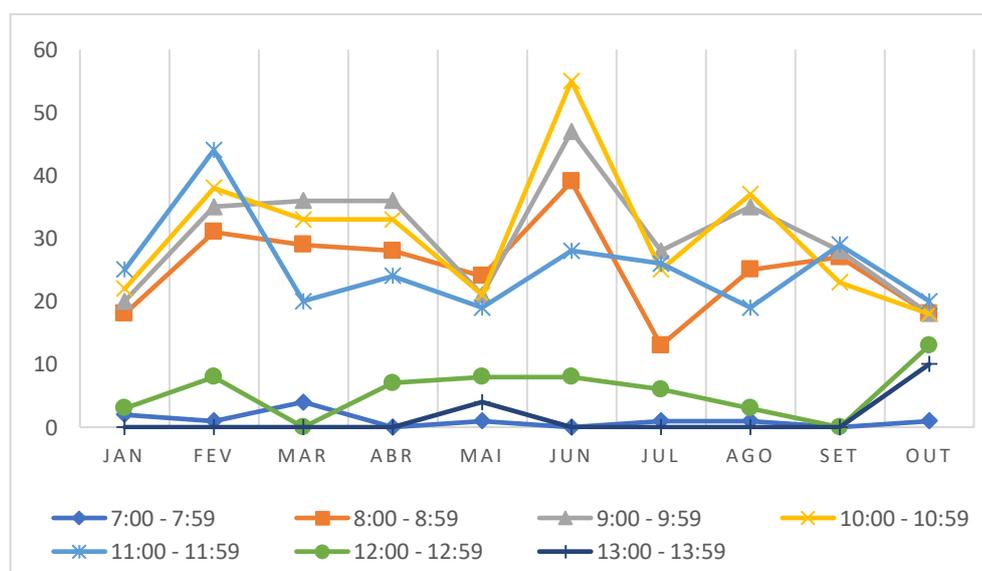
Período 01/01/2019 31/10/2019										
SÁBADO										
Horário	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT
7:00 - 7:59	2	1	4	0	1	0	1	1	0	1
8:00 - 8:59	18	31	29	28	24	39	13	25	27	18
9:00 - 9:59	20	35	36	36	21	47	28	35	28	18
10:00 - 10:59	22	38	33	33	21	55	25	37	23	18
11:00 - 11:59	25	44	20	24	19	28	26	19	29	20
12:00 - 12:59	3	8	0	7	8	8	6	3	0	13
13:00 - 13:59	0	0	0	0	4	0	0	0	0	10
TOTAL	90	157	122	128	98	177	99	120	107	98

Fonte: A Autora (2019).

Geralmente, o Banco de Sague realiza atendimento apenas na parte da manhã no sábado, mas de tempos em tempos é feito plantões e campanhas de doações onde se estende o horário de atendimento no sábado para que o Banco de Sangue possa receber mais possíveis doadores.

Pode-se ver os dados da Tabela 7 plotados na Figura 20 a seguir.

Figura 20: Quantidade de doares de hora em hora no sábado.



Fonte: A Autora (2019).

Mesmo tendo funcionamento diferenciado, é possível notar certo padrão na quantidade de doadores que comparecem ao banco de sangue aos sábados. É possível perceber que há maior movimento no meio da manhã, ou seja, entre 8h e 11h. Na primeira hora da manhã os doadores ainda estão começando a chegar no Banco de Sangue e é por esse motivo que a quantidade é baixa. Já nos horários de 12h a 13:59 a quantidade de doares é reduzida devida a proximidade com o horário de fechamento do mesmo.

5. Considerações Finais

Após a determinação das equações desejadas é possível concluir que o trabalho desenvolvido obteve sucesso, pois a metodologia empregada no estudo, a utilização do banco de dados e de um sistema operacional para a obtenção de distribuição de probabilidades foram eficazes.

Quanto aos objetivos específicos deste trabalho, pode-se concluir que não houve dificuldade em nenhuma etapa do desenvolvimento. O uso do banco de dados para a coleta e estratificação dos dados possibilitou a correta obtenção da quantidade de doadores que compareceram ao Banco de Sangue, ou seja, valores exatamente iguais aos ocorridos na realidade observada. O uso do método de Monte Carlo para a geração de distribuição de probabilidades por meio de um *software* que a autora já conhecia, o *ExpertFit®*, foi muito satisfatório. Pois, além da confiabilidade, apresentou facilidade na obtenção das distribuições estatísticas e tornou o progresso do estudo compreensível.

As equações matemáticas determinadas neste estudo são distribuição de probabilidades que possuem uma configuração que só é possível utilizar para a simulação de processos no *software FlexSim*. As funções encontradas, como por exemplo, Johnson SB, podem ser replicadas em outras plataformas, mas o uso dos parâmetros se molda apenas ao *software* empregado.

Os objetivos propostos para este trabalho foram alcançados de maneira concreta e pode-se então propor que em estudos futuros, a modelagem e simulação dos processos do Banco de Sangue analisados sejam executadas para que se chegue o mais próximo possível da realidade.

De maneira prática, este estudo contribui auxiliando no gerenciamento de funcionários do Banco de Sangue. Devido a seu caráter único, o departamento de recursos humanos analisa a realocação de funcionários de maneira pontual conforme a demanda de atendimento de

doadores cresce. Com as conclusões que este estudo proporcionou, é possível que o gerenciamento de funcionários seja feito de forma planejada e abrangente.

Referências

- ARENALES, M. ARMENTANO V. MORABITO, R. YANASSE, H. **Pesquisa Operacional: para cursos de engenharia**. 6. ed. Elsevier, RJ, 2007.
- BATEMAN, R. E. BOWDEN, R. O. GOGG, T. J. HARRELL, G. R. MOTT, J. R. A. MONTEVECHI, J. A. B. **Simulação de Sistemas: aprimorando processos de logística, serviços e manufatura**. 1ª Edição. Rio de Janeiro: Campus/Elsevier, 2013.
- BINDEL D., GOODMAN, J. **Principles of Scientific Computing**. New York University, New York, 2009.
- BUSSAB, W.O., MORETTIN, P.A. **Estatística Básica**. 7ª Edição. São Paulo: Editora Saraiva, 2011
- EXPERTFIT. **Averill M. Law & Associates**. Disponível em: <<http://www.averill-law.com/distribution-fitting/>>. Acesso em: 06/05/2019.
- LAMARCA, J. S. **Estudo do sistema de filas de uma casa lotérica**. Dissertação para obtenção do título de Mestre em Matemática da Universidade Federal do Triângulo Mineiro – UFTM, Departamento de Matemática. Uberaba, 2016.
- LUDWING, S. T. RODRIGUES, A. C. de M. **Doação de sangue: uma visão de marketing**. Cad. Saúde Pública, Rio de Janeiro, mai-jun, 2005.
- MENOSSE, Osvaldo Sérgio. **Estoque de segurança em refinarias de petróleo: análise por simulação**. Florianópolis, SC: UFSC. Programa de Pós-Graduação em Engenharia de Produção, 2002.
- MINISTÉRIO DA SAÚDE. Secretaria de Atenção à Saúde Departamento de Atenção Especializada e Temática. **Manual de Orientações para Promoção da Doação Voluntária de Sangue**. Brasília – DF, 2015.
- STURGES, H.A. The Choice of a Class Interval. *Journal of the American Statistical Association*. v. 21, n. 153, pp. 65–66, 2012
- OSWALT, R. M. **A review of blood donor motivation and recruitment**. *Transfusion* 1977; 17:123-35.
- PRADO, D. **Teoria das filas e da simulação**. 4. ed. Nova Lima: INDG Tecnologia e Serviços Ltda, 2009.
- PRODANOV, C. FREITAS, E. **Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico**. 2. Ed. Nova Hamburgo: Feevale, 2013.