

## Curso: Foundations

## Módulo: Probabilidades

### Aplicações

- 1) O IBGE divulgou em 2019 em seu livro Síntese de Indicadores Sociais: Uma Análise das Condições de Vida da População Brasileira – 2019<sup>(1)</sup>, os dados e informações relativas as condições de vida dos brasileiros. Um dos temas tratados foi o sobre a população brasileira com alguma ocupação com 14 anos de idade ou mais em algum trabalho principal. A tabela abaixo foi retirada da página 19 deste livro.

**Tabela 2 - População ocupada de 14 anos ou mais de idade no trabalho principal, segundo o tipo de ocupação - Brasil - 2012-2018**

Tipo de ocupação	População ocupada de 14 anos ou mais de idade						
	2012	2013	2014	2015	2016	2017	2018
<b>Valores absolutos (1 000 pessoas)</b>							
<b>Total</b>	<b>89 233</b>	<b>90 715</b>	<b>91 945</b>	<b>92 163</b>	<b>90 776</b>	<b>91 073</b>	<b>92 333</b>
Empregado com carteira	35 706	36 826	37 501	36 936	35 387	34 350	34 128
Empregado sem carteira	13 320	13 109	12 878	12 587	12 650	13 392	14 034
Trabalhador doméstico com carteira	1 886	1 819	1 914	1 987	1 975	1 864	1 738
Trabalhador doméstico sem carteira	4 303	4 249	4 083	4 119	4 184	4 319	4 495
Militar ou funcionário público	7 333	7 500	7 845	7 730	7 757	7 701	7 900
Conta própria	20 387	20 775	21 291	22 262	22 381	23 005	23 425
Empregador	3 541	3 728	3 811	4 020	4 245	4 215	4 484
Trabalhador familiar auxiliar	2 758	2 709	2 622	2 522	2 196	2 228	2 129
<b>Formal</b>	<b>52 212</b>	<b>54 153</b>	<b>56 040</b>	<b>56 245</b>	<b>55 366</b>	<b>53 986</b>	<b>54 039</b>
Empregado com carteira	35 706	36 826	37 501	36 936	35 387	34 350	34 128
Trabalhador doméstico com carteira	1 886	1 819	1 914	1 987	1 975	1 864	1 738
Militar ou funcionário público	7 333	7 500	7 845	7 730	7 757	7 701	7 900
Conta própria contribuinte	4 838	5 324	5 935	6 511	6 983	6 973	7 137
Empregador contribuinte	2 449	2 684	2 844	3 082	3 263	3 099	3 137
<b>Informal</b>	<b>37 021</b>	<b>36 562</b>	<b>35 906</b>	<b>35 918</b>	<b>35 410</b>	<b>37 087</b>	<b>38 294</b>
Empregado sem carteira	13 320	13 109	12 878	12 587	12 650	13 392	14 034
Trabalhador doméstico sem carteira	4 303	4 249	4 083	4 119	4 184	4 319	4 495
Conta própria não contribuinte	15 548	15 451	15 356	15 751	15 398	16 032	16 288
Empregador não contribuinte	1 092	1 044	967	939	982	1 116	1 347
Trabalhador familiar auxiliar	2 758	2 709	2 622	2 522	2 196	2 228	2 129

<sup>(1)</sup> Fonte: <https://biblioteca.ibge.gov.br/visualizacao/livros/liv101678.pdf>

Você foi convidado a realizar algumas análises envolvendo probabilidades com relação aos dados desta tabela. Considerando apenas o **ano de 2018**:

- a) Qual é o espaço amostral da população ocupada de 14 anos ou mais de idade?
- b) Qual são os dois principais eventos destacados em negrito na tabela?
- c) Qual a probabilidade de cada um dos dois principais eventos ocorrerem, separadamente, em relação ao espaço amostral?
- d) Qual a probabilidade de o evento “Conta própria não contribuinte” ocorrer em relação ao espaço amostral?
- e) Qual a probabilidade de o evento “Empregador contribuinte” ocorrer em relação:
  - I. Ao espaço amostral?
  - II. Ao grupo das pessoas ocupadas com 14 anos ou mais de idade que possuem um trabalho formal?

- f) Suponha que você veja a frase: “38,3 milhões de brasileiros tinham trabalhos informais em 2018, correspondendo a 71% dos trabalhadores formalizados”. Esta frase lhe causa algum incômodo? Por quê?
- 2) Você foi contratado para realizar alguns estudo preliminares com relação aos jogos da Mega Sena. A Principal questão aqui é levantar qual a probabilidade de algum apostador ter levado o prêmio nos jogos apresentados em “**sorteios\_megasena.xlsx**”. Assim:
- Preencha a coluna “Probabilidade de Alguém Ganhador” com a probabilidade de ao menos um jogador ganhar o concurso.  
**Dica:** o total de possibilidade de jogos com 6 números na Mega Sena é dado pela fórmula:  $COMBIN(60;6)$ , resultando num total de 50.063.860 jogos possíveis.
  - Preencha a coluna “Alguém Ganhador?” com 1, caso tenha havido um ganhador no concurso, e 0 caso contrário.
  - Vá na aba “Análises” e atualize a tabela de frequência. Calcule a probabilidade de algum ganhador utilizando as quantidades de concursos e quantidade de concursos com algum ganhador.
  - Compare a probabilidade calculada com a faixa de probabilidade na coluna A. Analise se está condizente e argumente o motivo.

- 3) Uma empresa mandou e-mails promocionais para sua carteira composta de 2.000 clientes. Como nem todos os clientes leem os informativos, a empresa continuou o envio de e-mails até que sua carteira de clientes inteira fosse impactada. Os resultados dessa ação seguem na tabela abaixo:

Quantidade de e-mails enviados até a 1ª leitura	1	2	3	4	5	6
Quantidade de clientes impactados	1050	300	250	200	160	40

- Qual a probabilidade de um destes clientes ser sorteado ao acaso ter feito a primeira leitura no primeiro e-mail recebido?
  - Qual a probabilidade de um destes clientes ser sorteado ao acaso ter recebido mais do que 3 e-mails até a primeira leitura?
  - Crie na tabela uma linha com a quantidade de e-mails enviados total (qtd e-mails enviados \* qtd de clientes).
  - Analisando os itens anteriores, o que você achou dessa ação da empresa? Faria alguma modificação na estratégia?
- 4) Segundo a organização da Megaseña, todos os 60 números têm a mesma chance de ocorrência. Em linguagem estatística, o número  $Y$  (número sorteado da Megaseña) é uma variável aleatória que segue a distribuição **Uniforme**. Ou seja,  $Y \sim U(1;60)$ . Para verificar se a organização fala a verdade, siga os passos abaixo:
- Abra o arquivo “**sorteios\_megasena.xlsx**”
  - Preencha a seguinte tabela usando as informações do enunciado e do arquivo indicado no item a)

Número da Megasena	Frequência Observada	Frequência Esperada (segundo modelo Uniforme)
1		
2		
...		
60		

**Responda:**

- A. O modelo Uniforme é adequado para o conjunto de dados da Megasena? Justifique.  
 B. Você apostaria na sequência 01, 02, 03, 04, 05 e 06? Justifique.

- 5) Uma empresa realizou uma avaliação demográfica de seus funcionários, através de uma amostra de 500 pessoas. Uma das variáveis analisadas foi a idade, identificando que ela se aproxima a uma distribuição Normal com média 30 e variância 16. A maneira que podemos escrever isso em notação é:  $X \sim N(30; 16)$ , sendo  $X$  a idade dos funcionários. Outras notações importantes são:  $P(X \leq i)$  - probabilidade da idade ser menor ou igual que  $i$ , e  $P(X \geq i)$  - a probabilidade da idade ser maior ou igual que  $i$ . Dessa forma, responda:
- $P(X \geq 32)$  – Qual a probabilidade de um funcionário ter 32 anos ou mais?
  - $P(X \leq 30)$  – Qual a probabilidade de um funcionário ter até a média de idade, ou seja, menor ou igual a 30 anos?
  - $P(28 \leq X \leq 32)$  – Qual a probabilidade de um funcionário ter entre 28 e 32 anos?
  - $P(20 \leq X \leq 24)$  – Qual a probabilidade de um funcionário ter entre 20 e 24 anos?
  - $P(36 \leq X \leq 40)$  – Qual a probabilidade de um funcionário ter entre 36 e 40 anos?
  - Analisando os itens c) e d), responda: por que as probabilidades são diferentes apesar da amplitude do intervalo ser igual? Faça a mesma análise para o item d) e e): as probabilidades são iguais ou diferentes? Por quê?

**Exercícios sobre Teoria Axiomática, Distribuição Binomial e Poisson (Opcionais)**

- 6) Um banco de investimentos tem 10 mil clientes dos quais 4 mil são clientes Private. Cerca de 500 clientes têm perfil de risco Arrojado, 700 têm perfil de risco Conservador, 100 são do Private e Arrojados e 200 são Private e Conservadores. Ao escolher um cliente ao acaso, pergunta-se qual a probabilidade de:
- Ser um cliente Private
  - Ser Private e Conservador
  - Não ter perfil de risco Arrojado ou Conservador
  - Ser Private ou ter perfil de risco entre Arrojado e Conservador
  - Não ser Private e nem ter perfil de risco entre Arrojado e Conservador
- 7) Uma central telefônica recebe em média 1.000 chamadas por hora, e sabe-se que cerca de 2% dessas chamadas são para cancelamento. Como o cancelamento é um assunto delicado, existe uma equipe

dedicada para convencer os clientes a desistirem do cancelamento e manterem o serviço contratado. Para poder dimensionar adequadamente essa equipe, realize as seguintes análises:

- a) Qual é a probabilidade de ocorrência de mais de 20 ligações sobre cancelamento nessas 1.000 chamadas em 1 hora?

**Dica:** utilize a função DISTR.BINOM()

- b) Se cada atendente demora em média 6 minutos por cliente, quantos atendentes são necessários para que a probabilidade de nenhuma ligação ficar em espera seja inferior a 5%?

**Dica:** utilize a função INV.BINOM() para calcular a quantidade de ligações de cancelamento

- 8) Em um aeroporto aterrissam, em média, 10 aviões por hora. Qual a probabilidade de que:

- a) Em uma hora aterrissem no máximo 4 aviões?  
b) Em duas horas aterrissem ao menos 15 aviões?

**Dica:** Utilize a função DIST.POISSON() do Excel.

- 9) Um motorista do Uber deseja modelar o atingimento de sua meta semanal de corridas. Para isso, ele categorizou um período de uma semana (56h de trabalho) pela quantidade de corridas por hora. Os resultados seguem na tabela abaixo:

Corridas/hora	0	1	2	3	4	5	6	>= 7
Qtd horas (das 56hrs)	2	4	5	7	15	10	8	5

O motorista, após fazer um curso de Analytics na [Preditiva.ai](#), decide verificar se a distribuição de Poisson se ajusta bem aos dados. Para isso, pergunta-se:

- a) Qual a taxa de ocorrência  $\lambda$  da Poisson que mais se ajusta aos dados do motorista?  
b) Supondo que o motorista ganhe em cada corrida R\$ 12 reais e que sua meta semanal (com 8h de diárias) seja R\$ 1.000 reais, qual a probabilidade de que ele bata a meta? Utilize uma Poisson com a taxa de ocorrência do item anterior.