# Estatística Avançada - Aula 02

## Variáveis Aleatórias e Distribuição de Probabilidades II

Kaique Matias de Andrade Roberto

Ciências Atuariais - Ciências Econômicas

HECSA - Escola de Negócios

FIAM-FAAM-FMU

# Conteúdo

# Conceitos que aprendemos em Aulas anteriores

# Conceitos que aprendemos em Aulas anteriores

- variáveis aleatórias;
- variáveis aleatórias discretas;
- distribuições uniforme discreta, Bernoulli, binomial e Poisson;
- condições de convergência de distribuições binomiais para uma Poisson.

# Variável Aleatória

# Variável Aleatória

Vamos retomar os conceitos e terminologias relacionados à Teoria das Probabilidades que vimos na Aula 01.

**Definição 2.1**

Um **experimento** consiste em qualquer processo de observação ou medida. Um **experimento aleatório** é aquele que gera resultados imprevisíveis, de modo que, se o processo for repetido inúmeras vezes, torna-se impossível prever seu resultado.

# Variável Aleatória

**Definição 2.2**

O espaço amostral $S$ consiste em todos os possíveis resultados de um experimento.

# Variável Aleatória

**Definição 2.3**

Um **evento** é qualquer subconjunto de um espaço amostral.

**Definição 2.4 (Variável Aleatória)**

Consideremos $\mathcal{E}$ um experimento aleatório e $S$ o espaço amostral associado ao experimento. A função $X$ que associa a cada elemento $s \in S$ um número real $X(s)$ é denominada **variável aleatória**.

# Variável Aleatória

As variáveis aleatórias podem ser discretas ou contínuas.

# Variável Aleatória

**Exercício 2.1**

Considere o experimento de lançar uma certa moeda e observar se ocorre cara (k) ou coroa (c). Descreva o comportamento da variável

$$N = \text{Número de caras em dois lançamentos dessa moeda.}$$

Em seguida, calcular $E[X]$, $\text{Var}(X)$ e a distribuição acumulada de $X$.

# Distribuição Geométrica e Hipergeométrica

A **distribuição geométrica**, assim como a binomial, considera sucessivos ensaios de Bernoulli independentes, todos com probabilidade de sucesso $p$.

Porém, em vez de utilizar um número fixo de tentativas, elas serão realizadas até que o primeiro sucesso seja obtido.

# Distribuição Geométrica e Hipergeométrica

A distribuição geométrica apresenta duas parametrizações ("fórmulas") distintas.

A primeira parametrização considera sucessivos ensaios de Bernoulli independentes, com probabilidade de sucesso $p$ em cada ensaio, até que ocorra um sucesso. Nesse caso, não podemos incluir o zero como um possível resultado, de modo que o domínio é suportado pelo conjunto $\{1, 2, 3, ...\}$.

Por exemplo, podemos considerar a quantidade de lançamentos de uma moeda até a primeira cara, a quantidade de peças produzidas até a próxima defeituosa, etc.

A segunda parametrização conta o número de falhas ou fracassos antes do primeiro sucesso. Como aqui é possível obter sucesso já no primeiro ensaio de Bernoulli, incluímos o zero como resultado possível, de modo que o domínio é suportado pelo conjunto $\mathbb{N} = \{0, 1, 2, 3, ...\}$.

**Definição 3.1**

Seja $X$ a variável aleatória que representa o número de tentativas até o primeiro sucesso.A variável X tem distribuição geométrica com parâmetro $p$, denotada por $X \sim \text{Geo}(p)$, se sua função de probabilidade for dada por:

$$P(X = x) = p(1-p)^{x-1},\ x \geq 1.$$

Para o segundo caso, consideremos $Y$ a variável aleatória que representa o número de falhas ou fracassos antes do primeiro sucesso.A variável $Y$ tem distribuição geométrica com parâmetro $p$, denotada por $Y \sim \text{Geo}(p)$, se sua função de probabilidade for dada por:

$$P(Y = y) = p(1 - p)^y, \; y \geq 0.$$

Em ambos os casos, a sequência de probabilidades é uma progressão geométrica.

Pode-se demonstrar que, para $X$ e $Y$ como acima valem

$$E(X) = \frac{1}{p}, \text{ e } \mathrm{Var}(X) = \frac{1-p}{p};$$

e

$$E(Y) = \frac{1}{p}, \text{ e } \mathrm{Var}(Y) = \frac{1-p}{p}.$$

**Exemplo 3.2**

Uma empresa fabrica determinado componente eletrônico, de modo que, ao final do processo, cada componente é testado, um a um. Suponha que a probabilidade de um componente eletrônico estar defeituoso seja de $0,05$. Determine a probabilidade de que o primeiro defeito seja encontrado no oitavo componente testado. Calcule também o valor esperado e a variância da variável aleatória.

A distribuição geométrica é a única distribuição discreta que tem a propriedade da falta de memória (no caso das distribuições contínuas, veremos que a distribuição exponencial também apresenta essa propriedade).

Isso significa que, se um experimento for repetido antes do primeiro sucesso, então, dado que o primeiro sucesso ainda não ocorreu, a função de distribuição condicional do número de tentativas adicionais não depende do número de fracassos ocorridos até então.

# Distribuição Geométrica e Hipergeométrica

Assim, para quaisquer dois inteiros positivos $s$ e $t$, se $X$ for maior do que $s$, então a probabilidade de que $X$ seja maior do que $s + t$ é igual à probabilidade incondicional de $X$ ser maior do que $t$:

$$P(X > s + t | X > s) = P(X > t).$$

A distribuição hipergeométrica também está relacionada com um experimento de Bernoulli.

# Distribuição Geométrica e Hipergeométrica

Porém, diferentemente da amostragem binomial, em que a probabilidade de sucesso é constante, na distribuição hipergeométrica, como a amostragem é sem reposição, à medida que os elementos são retirados da população para formar a amostra, o tamanho da população diminui, fazendo com que a probabilidade de sucesso varie.

A distribuição hipergeométrica descreve o número de sucessos na amostra de *n* elementos, extraída de uma população finita sem reposição.

Por exemplo, consideremos uma população com $N$ elementos, dos quais $M$ possuem determinado atributo. A distribuição hipergeométrica descreve a probabilidade de que, em uma amostra com $n$ elementos distintos extraídos aleatoriamente da população sem reposição, exatamente $k$ possuem tal atributo ($k$ sucessos e $n - k$ fracassos).

**Definição 3.3**

Seja $X$ uma variável aleatória que representa o número de sucessos obtidos a partir dos n elementos retirados da amostra. A variável $X$ segue distribuição hipergeométrica com parâmetros $N$, $M$, $n$, denotada por $X \sim \text{Hip}(N, M, n)$, se sua função de probabilidade for dada por:

$$P(X = k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}, \; 0 \leq k \leq \min(M, n).$$

Se $X \sim \text{Hip}(N, M, n)$, pode-se demonstrar que

$$E(X) = \frac{nM}{N} \text{ e } \text{Var}(X) = \frac{nM(N-M)(N-n)}{N^2(N-1)}.$$

**Exemplo 3.4**

Uma urna contém 15 bolas, das quais 5 delas são vermelhas. São escolhidas 7 bolas ao acaso, sem reposição. Determine:

a - A probabilidade de que exatamente duas bolas vermelhas sejam sorteadas.

b - A probabilidade de que pelo menos duas bolas vermelhas sejam sorteadas.

c - O número esperado de bolas vermelhas sorteadas.

d - A variância do número de bolas vermelhas sorteadas.

# Variável Aleatória Contínua

**Definição 4.1**

Uma **variável aleatória contínua** é aquela que pode assumir diversos valores num intervalo de números reais.

# Variável Aleatória Contínua

**Exemplo 4.2**

Como exemplos de variáveis aleatórias contínuas, podemos citar a renda familiar, o faturamento da empresa ou a altura de determinada criança.

# Variável Aleatória Contínua

Para calcular probabilidades no contexto das variáveis contínuas vamos recorrer ao Cálculo Diferencial e Integral.

# Variável Aleatória Contínua

**Definição 4.3**

Uma variável aleatória contínua $X$ está associada a uma função $f(x)$, denominada função **densidade de probabilidade (f.d.p.)** de $X$, que satisfaz a seguinte condição:

$$\int_{-\infty}^{\infty} f(x)\, dx = 1,\ f(x) \geq 0.$$

# Variável Aleatória Contínua

Para quaisquer $a$ e $b$, tal que $-\infty < a < b < \infty$, a probabilidade de que a variável aleatória $X$ assuma valores nesse intervalo é:

$$P(a \leq X \leq b) = \int_a^b f(x)\,dx.$$

# Variável Aleatória Contínua

A esperança matemática (valor esperado ou médio) de uma variável aleatória contínua $X$ com função densidade de probabilidade $f(x)$ é dada pela expressão:

$$E(X) = \int_{-\infty}^{\infty} xf(x)\,dx.$$

# Variável Aleatória Contínua

A variância de uma variável aleatória contínua $X$ com função densidade de probabilidade $f(x)$ é calculada como:

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} (x - E(x))^2 f(x)\, dx.$$

# Variável Aleatória Contínua

Como no caso de variáveis aleatórias discretas, podemos calcular probabilidades associadas a uma variável aleatória contínua X a partir de uma função de distribuição acumulada.

# Variável Aleatória Contínua

A função de **distribuição acumulada** $F(x)$ de uma variável aleatória contínua X com função densidade de probabilidade $f(x)$ é definida por:

$$F(x) = P(X \leq x), -\infty < x < \infty.$$

# Variável Aleatória Contínua

Agora vamos conhecer algumas dentre as principais distribuições contínuas.

# Distribuição Uniforme Contínua

# Distribuição Uniforme Contínua

A distribuição uniforme é a mais simples para variáveis aleatórias contínuas, sendo utilizada para modelar a ocorrência de eventos cuja probabilidade é constante em intervalos de mesma amplitude.

# Distribuição Uniforme Contínua

**Definição 5.1**

Uma variável aleatória $X$ tem **distribuição uniforme** no intervalo $[a, b]$, denotada por $X \sim U[a, b]$, se sua função densidade de probabilidade for dada por:

$$f(x) = \begin{cases} \dfrac{1}{b-a}, & \text{se } a \leq x \leq b \\ 0 \text{ caso contrário.} \end{cases}$$

# Distribuição Uniforme Contínua

Pode-se demonstrar que se $X \sim U[a, b]$ então

$$E(X) = \frac{a+b}{2}, \text{ e } \mathrm{Var}(X) = \frac{(b-a)^2}{2}.$$

Já a função de distribuição acumulada de $X$ é

$$f(x) = \begin{cases} 0, \text{ se } x < a, \\ \dfrac{x-a}{b-a}, \text{ se } a \leq x < b \\ 0 \text{ se } x \leq b. \end{cases}$$

**Exemplo 5.2**

A variável aleatória $X$ representa o tempo de utilização dos caixas eletrônicos de um banco (em minutos) e segue uma distribuição uniforme no intervalo $[1, 5]$. Determine:

a - $P(X < 2)$;

b - $P(X > 3)$;

c - $P(3 < X < 4)$;

d - $E(X)$;

e - $\text{Var}(X)$.

# Distribuição Exponencial

# Distribuição Exponencial

A distribuição exponencial tem como principal característica a propriedade de não possuir memória, isto é, o tempo de vida futuro ($t$) de determinado objeto tem a mesma distribuição, independente do seu tempo de vida passada ($s$), para quaisquer $s, t > O$:

$$P(X > s + t | X > s) = P(X > t).$$

# Distribuição Exponencial

**Definição 6.1**

Uma variável aleatória contínua $X$ tem **distribuição exponencial** com parâmetro $\lambda > O$, denotada por $X \sim \exp(\lambda)$, se sua função densidade de probabilidade for dada por:

$$f(X) = \begin{cases} \lambda e^{-\lambda x}, \text{ se } x \geq 0 \\ 0 \text{ se } x < 0. \end{cases}$$

# Distribuição Exponencial

Pode-se demonstrar que se $X \sim \exp(\lambda)$ então

$$E(X) = \frac{1}{\lambda} \text{ e } \mathrm{Var}(X) = \frac{1}{\lambda^2};$$

e a função de distribuição acumulada de $X$ é dada por

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, \text{ se } x \geq 0 \\ 0 \text{ se } x < 0. \end{cases}$$

Logo, $P(X > x) = e^{-\lambda x}$.

**Exemplo 6.2**

Uma indústria fabrica lâmpadas especiais que ficam em operação continuamente. A empresa oferece a seus clientes a garantia de reposição, caso a lâmpada dure menos de 50 horas. A vida útil dessas lâmpadas é modelada através da distribuição Exponencial com parâmetro $1/8000$. Determine a proporção de trocas por defeito de fabricação.

# Outras Distribuições Contínuas

# Outras Distribuições Contínuas

Existem inúmeras distribuições contínuas. Para nós serão importantes as distribuições:

- Normal;
- $t$ de Student;
- Qui-quadrado.

# Outras Distribuições Contínuas

As três distribuições anteriores tem relação com a **distribuição Gamma** (que não vamos estudar neste curso mas é extremamente importante em várias situações da área de vocês).

# Comentários Finais

# Comentários Finais

Em resumo, na aula de hoje nós:

- variáveis aleatórias;
- distribuições geométrica e hipergeométrica;
- variáveis aleatórias contínuas;
- distribuições uniforme contínua e exponencial;
- outros exemplos de distribuições contínuas.

# Comentários Finais

Nas próxima aula nós vamos focar na Distribuição Normal. Mais especificamente vamos:

- definir o que é a Distribuição Normal;
- dar Exemplos de situações que envolvem normalidade;
- entender os $z$-scores;
- calcular as principais probabilidades envolvendo a Distribuição Normal $N(0, 1)$;
- extrapolar esses cálculos para uma Normal $N(\mu, \sigma^2)$ qualquer.

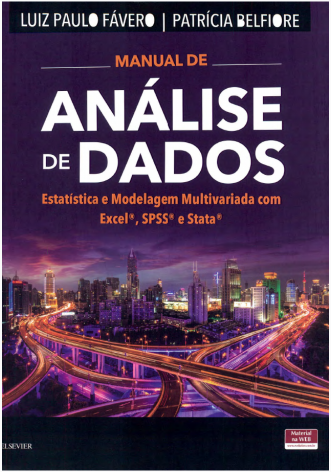
## Comentários Finais

**ATIVIDADE PARA ENTREGAR (E COMPOR A NOTA N1)**

Em grupos de até 4 integrantes resolva três dentre os Exercícios 2.1-2.6.

# Referências

# Referências



# Bons Estudos!