Dev - Probabilidade I

Síntese

Limites de Conjuntos

▼

- A definição de conceitos de convergência para variáveis aleatórias baseia-se em manipulações de sequências de eventos que requerem limites de conjuntos. Seja $A_n \subset \Omega$ definimos

$$\inf_{k \geq n} A_k := \bigcap_{k=n}^{\infty} A_k, \quad \sup_{k \geq n} A_k := \bigcup_{k=n}^{\infty} A_k$$

$$\liminf_{n \to \infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k,$$

$$\limsup_{n \to \infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k.$$

- O limite de uma sequência de conjuntos é definido da seguinte forma: Se para alguma sequência $\{B_n\}$ de subconjuntos

$$\limsup_{n \to \infty} B_n = \liminf_{n \to \infty} B_n = B,$$

então B é chamado de limite de B_n e escrevemos $\lim_{n\to\infty} B_n = B$ ou $B_n \to B$ será demonstrado em breve que

$$\liminf_{n \to \infty} A_n = \lim_{n \to \infty} \left(\inf_{k \geq n} A_k \right)$$

and

$$\limsup_{n \to \infty} A_n = \lim_{n \to \infty} \left(\sup_{k \geq n} A_k \right).$$

- **Exemplo:**

$$\liminf_{n\to\infty}[0, n/(n+1)) = \limsup_{n\to\infty}[0, n/(n+1)) = [0, 1). \qquad \square$$

Podemos agora dar uma interpretação de $\liminf_{n\to\infty} A_n$ e $\limsup_n \to \infty A_n$.

LEMA: Seja $\{A_n\}$ uma sequência de subconjuntos de Ω.

a. Para $\limsup$ temos a interpretação

$$\begin{aligned}\limsup_{n\to\infty} A_n &= \left\{\omega : \sum_{n=1}^{\infty} 1_{A_n}(\omega) = \infty\right\} \\ &= \{\omega : \omega \in A_{n_k}, k = 1, 2 \dots\}\end{aligned}$$

para alguma subsequência n_k dependendo de ω. Consequentemente, escrevemos

$$\limsup_{n\to\infty} A_n = [A_n \ i.o.\ ]$$

onde $i.o.$ significa infinitamente frequentemente.

b. Para $\liminf$ temos a interpretação

$$\begin{aligned}\liminf_{n\to\infty} A_n &= \{\omega : \omega \in A_n \textit{ for all } n \textit{ except a finite number }\} \\ &= \{\omega : \sum_n 1_{A_n^c}(\omega) < \infty\} \\ &= \{\omega : \omega \in A_n, \forall n \geq n_0(\omega)\}.\end{aligned}$$

Prova (a): Se

$$\omega \in \limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k,$$

então para todo n, $\omega \in \cup_{k\geq n} A_k$ e assim para todo n, existe algum $k_n \geq n$ tal que $\omega \in A_{k_n}$, e portanto

$$\sum_{j=1}^{\infty} 1_{A_j}(\omega) \geq \sum_n 1_{A_{k_n}}(\omega) = \infty,$$

que implica

$$\omega \in \left\{\omega : \sum_{n=1}^{\infty} 1_{A_n}(\omega) = \infty\right\};$$

portanto

$$\limsup_{n\to\infty} A_n \subset \{\omega : \sum_{j=1}^{\infty} 1_{A_j}(\omega) = \infty\}.$$

Inversamente, se

$$\omega \in \{\omega : \sum_{j=1}^{\infty} 1_{A_j}(\omega) = \infty\},$$

então existe $k_n \to \infty$ tal que $\omega \in A_{k_n}$, e portanto para todo n, $\omega \in \cup_{j \geq n} A_j$ tal que $\omega \in \limsup_{n\to\infty} A_n$. Por definição

$$\{\omega : \sum_{j=1}^{\infty} 1_{A_j}(\omega) = \infty\} \subset \limsup_{n\to\infty} A_n.$$

Isso prova a inclusão do conjunto em ambas as direções e mostra a igualdade.

Prova (b): A prova de (b) é análoga.

Leis de De Morgan

- $(A \cup B)^c = A^c \cap B^c$
- $(A \cap B)^c = A^c \cup B^c$

Propriedades distributiva

- $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$
- $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$

Arranjos

$$\boxed{A_{m,r} = \frac{m!}{(m-r)!}}$$

- Seja M um conjunto com m elementos, isto é, $M = \{a_1, a_2, ..., a_m\}$. Chamamos de arranjo dos m elementos tomados r a r $(1 \leqslant r \leqslant m)$ a qualquer r-upla (sequência de r elementos) formada com elementos de M, todos distintos.

Arranjos com repetição

$$\boxed{(AR)_{m,r} = \underbrace{m \cdot m \cdot ... \cdot m}_{r \text{ vezes}} = m^r}$$

- Seja M um conjunto com m elementos, isto é, $M = \{a_1, a_2, ..., a_m\}$. Chamamos arranjo com repetição dos m elementos, tomados r a r, toda r-upla ordenada (sequência de tamanho r) formada com elementos de M não necessariamente distintos.

Permutações

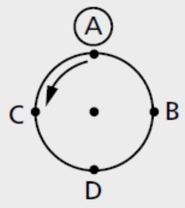
$$\boxed{P_m = m \cdot (m-1) \cdot (m-2) \cdot ... \cdot 3 \cdot 2 \cdot 1}$$

- Seja M um conjunto com m elementos, isto é, $M = \{a_1, a_2, ..., a_m\}$. Chamamos de permutação dos m elementos a todo arranjo em que $r = m$.

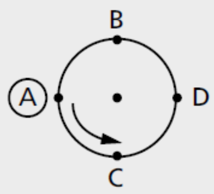
Permutações Circular

- Quando elementos são dispostos ao redor de um círculo, a cada disposição possível chamamos permutação circular.
- Duas permutações circulares são consideradas idênticas, quando:

1) Tomando o elemento A, a sequência encontrada é (A,C, D, B).



2) Tomando o elemento A, a sequência encontrada é (A, C, D, B).



-

$$\boxed{x = \frac{n!}{n} = (n-1)!}$$

Permutações Com Repetição

$$P_n^{n_1,\, n_2,\, \ldots,\, n_r} = \frac{n!}{n_1!\; n_2! \ldots n_r!}$$

Combinações

$$C_{m,\,r} = \binom{m}{r} = \frac{m!}{r!\,(m-r)!} \qquad \forall m, r \in \mathbb{N}^*, r < m$$

- **Casos particulares:**

$$1^{\underline{o}} \text{ caso: } m, r \in \mathbb{N}^* \text{ e } r = m$$

$$\begin{cases} C_{m,\,m} = 1 \\ \dfrac{m!}{m!(m-m)!} = 1 \end{cases}$$

$$2^{\underline{o}} \text{ caso: } m \in \mathbb{N}^* \text{ e } r = 0$$

$$\begin{cases} C_{m,\,0} = 1 \text{ (o único subconjunto com 0 elemento é o vazio)} \\ \dfrac{m!}{0!\,(m-0)!} = 1 \end{cases}$$

3º caso: m = 0 e r = 0

$$\begin{cases} C_{0,0} = 1 \text{ (o único subconjunto do conjunto vazio é o próprio vazio)} \\ \dfrac{0!}{0!\,(0-0)!} = 1 \end{cases}$$

Teorema Conjuntos

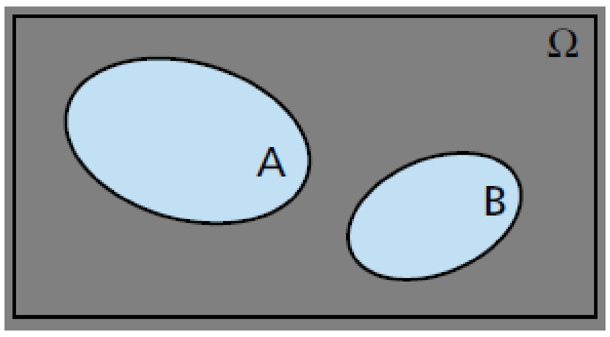
- Se $A \subset B$, então $P(A) \leqslant P(B)$.
- Se A é um evento, então $0 \leqslant P(A) \leqslant 1$.
- $P(A) = 1 - P(A^c)$
- Sendo A e B dois eventos quaisquer, vale $P(B) = P(B \cap A) + P(B \cap A^c)$.
- Regra da Adição de Probabilidades. Se A e B são eventos, então $P(A \cup B) = P(A) + P(B) – P(A \cap B)$
- Para eventos quaisquer $A_1, A_2, \ldots$

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} P(A_i)$$

- Se os eventos $A_1, A_2, \ldots$ são mutuamente exclusivos, então:

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

- Se A e B são mutuamente exclusivos $(A \cap B = \emptyset)$, então $P(A \cup B) = P(A) + P(B) – P(\emptyset) = P(A) + P(B)$.



- Se A é um evento, então $P(A^C) = 1 – P(A)$
- **Continuidade da Probabilidade:** Se $A_n \uparrow A$ então $P(A_n) \uparrow P(A)$. De forma similar, se $A_n \downarrow A$ então $P(A_n) \downarrow P(A)$. A notação $A_n \uparrow A$ indica que temos uma sequência monótona não decrescente ($n \leq n+1$) de eventos $A_i, A_2, \ldots$ tais que $\lim_{i \to \infty} A_i = A = \bigcup_{i=1}^{\infty} A_i$.

Probabilidade condicional

$$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)}} \quad P(B) > 0$$

Teorema da multiplicação

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \Rightarrow \boxed{P(A \cap B) = P(B) \cdot P(A|B)}$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \Rightarrow \boxed{P(A \cap B) = P(A) \cdot P(B|A)}$$

Independência de dois eventos

- Dois eventos A e B de um espaço amostral Ω, são chamados **independentes** se $P(A \cap B) = P(A) \cdot P(B)$. Ou seja, se os eventos são independentes a ocorrência de um não altera a probabilidade de ocorrência do outro.

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(B) \cdot P(A|B)}{P(A)} = \frac{P(B) \cdot P(A)}{P(A)} = P(B)$$

  Dois eventos A e B são independentes se, e somente se a definição é verdadeira:

$$P(A \cap B) = P(A)P(B)$$

- **Obs:** Sejam A e B eventos de um espaço amostral Ω tais que $P(A) > 0$ e $P(B) > 0$. Se A e B são **INDEPENDENTES**, então A e B não podem ser **MUTUAMENTE EXCLUDENTES**.

Variáveis Aleatórias (v.a)

▼ **Definição**

- A variável aleatória vai associar um valor real para cada valor do espaço amostral Ω.

  Ex: Lançar duas moedas consecutivamente

  - Notação $C = cara$, $K = coroa$.
  - $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$
  - $\omega_1 = (C, C)$; $\omega_2 = (C, K)$; $\omega_3 = (K, C)$; $\omega_4 = (K, K)$
  - Defina X: O número de coroas observadas.

$$\begin{aligned} \omega_1 \rightarrow X &= 0 \\ \omega_2 \,\text{ou}\, \omega_3 \rightarrow X &= 1 \\ \omega_4 \rightarrow X &= 2 \end{aligned}$$

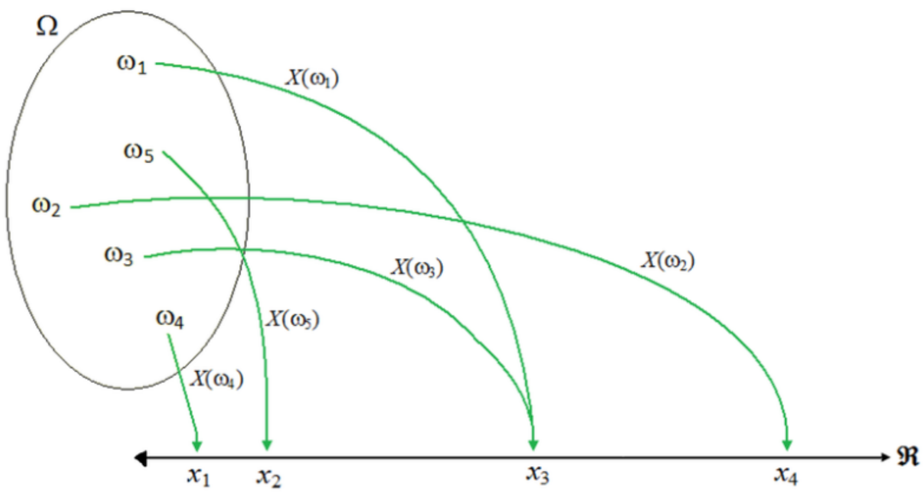
- De uma maneira mais formal: Uma variável aleatória é uma função definida num espaço amostral, que assume valores reais.
- Em outras palavras. Uma variável aleatória X representa um valor real, associado a cada resultado de um experimento de probabilidade. Ou seja para cada valor do Ω vai existir um valor x dentro da variável aleatória X.
- **Obs:** As variáveis aleatórias são representadas por letras maiúsculas X. E os valores assumidos pelas variáveis aleatórias são representados por letras minúsculas (x).

Definição

Sejam ε um experimento aleatório e Ω o espaço amostral associado a este. Uma *função* X, que associe a cada elemento $\omega \in \Omega$ um número real, $X(\omega)$, é denominada variável aleatória (v.a).

Notação:

- Letras maiúsculas (variável aleatória), X, Y, Z, etc.
- Letras minúsculas, (valor assumido pela variável aleatória), x, y, z etc.



Exemplos:

(a) Número de peças defeituosas entre n peças retiradas de uma linha de produção.
$X = \{x|x \in \{0, 1, 2, \ldots, D\}\}$ onde D é o total de peças defeituosas

(b) Número de veículos que passam por um posto de pedágio durante uma hora.
$X = \{x|x \in \{0, 1, 2, \ldots\}\}$

(c) Sexo de um indivíduo selecionado ao acaso de uma população
$X = \{x|x = 0 \text{ ou } x = 1\}$

(d) Tempo de duração de um componente de um circuito.
$X = \{x|x \in [0, \infty)\}$

(e) Peso de animais sujeitos a uma dieta de engorda.
$X = \{x|x \in [0, \infty)\}$

- **Definição:** As variáveis aleatórias que assumem valores em um conjunto enumerável serão denominadas **Discretas** e aquelas que assumem valores em um intervalo da reta real serão denominadas **Contínuas**.

▼ **Variáveis Aleatórias Discreta**

- Uma variável aleatória é discreta se ela assume um **número finito** de valores ou assume um **número infinito de valores numeráveis** (contáveis).
- Ou seja, podemos dizer que uma variável é discreta quando seus valores puderem ser listados.

Por exemplo: o número de ligações recebidas por dia em um escritório pode ser um valor igual a 0, 1, 2, 3, 4, ... Assim, definimos a variável aleatória X:

X: número de ligações recebidas pelo escritório.

Os valores que essa variável pode assumir são x=0, 1, 2, 3, ... Dessa forma, se escrevermos X=3 estamos dizendo que “o número de ligações recebidas pelo escritório (X) é igual a 3 ligações (x)”.

▼ Variáveis Aleatórias Continua

- Uma variável aleatória é contínua se ela possui um **número incontável** de possíveis resultados.
- Ou seja, uma variável é dita contínua quando os valores que ela pode assumir puderem ser representados como um **intervalo na reta dos números reais**.
- Neste caso, os valores assumidos por uma variável contínua **não podem ser listados**, visto que **são infinitos** os possíveis valores dessa variável.

Por exemplo: consideremos o tempo de duração de uma ligação recebida em minutos (incluindo frações de minutos). Neste caso, podemos definir uma variável aleatória Y da seguinte forma:

Y: tempo de duração de uma ligação em minutos.

Perceba que os valores de Y podem assumir qualquer valor em um intervalo real. Suponhamos, para facilitar, que o tempo máximo de uma ligação seja de 120 minutos. Neste caso, os valores y pertencem ao intervalo [0, 120].

Distribuição de Probabilidade

▼ Função Discreta de Probabilidade

- Para cada valor $x_1, x_2, x_3...$ de uma variável aleatória discreta (X) pode-se determinar uma probabilidade correspondente a esse valor. Ou seja, para cada x_i vai ser atribuído uma probabilidade de ocorrer.
- Representamos essa probabilidade da seguinte maneira: $P(x_i)$ que é a mesma coisa que $\mathbb{P}(X = x_i)$. Essas duas expressões são chamadas de **Função de Probabilidade** ou **Funções Discretas de Probabilidade**.

$$P(x_i) = \mathbb{P}(X = x_i)$$

▼ Definição de Distribuição de Probabilidade

- A **Distribuição de probabilidade** vai ser a coleção de todas as probabilidades que foram atribuídas aos x_i.
- Uma distribuição de probabilidades deve satisfazer a duas condições:
    i. A probabilidade de cada valor $p(x_i)$ da variável aleatória tem que ser um número de 0 á 1.

$$0 \leq \mathbb{P}(X = x) \leq 1 \qquad ou \qquad 0 \leq P(x) \leq 1$$

    ii. A soma de todas as probabilidades tem que ser igual a 1.

$$\sum_i \mathbb{P}(X = x_i) = 1 \qquad ou \qquad \sum_i p(x_i) = 1$$

Definição

Seja X uma variável aleatória discreta. Portanto, X assume no máximo um número infinito enumerável de valores $x_1, x_2, \ldots$. A cada possível resultado x_i associaremos um número $p(x_i) = \mathbb{P}(X = x_i)$, denominado probabilidade de x_i satisfazendo as seguintes condições:

(a) $p(x_i) \geq 0$ para todo i,

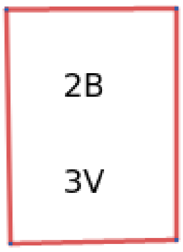
(b) $\sum_i p(x_i) = 1$.

A função $p(x_i)$ é chamada *função de probabilidade* e a coleção de pares $[x_i, p(x_i)]$, $\quad i = 1, 2, \ldots$, é denominada *distribuição de probabilidade* de X.

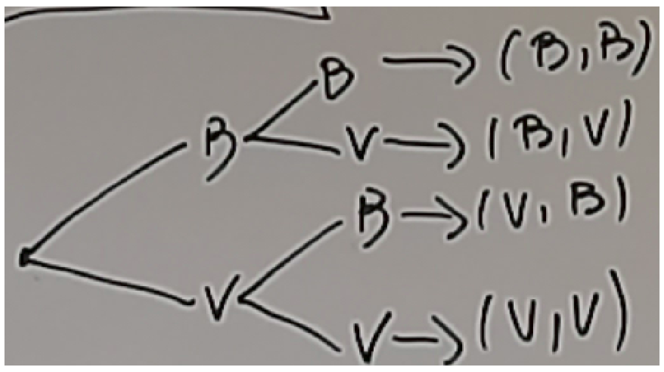
Exemplos

Ex 1

- Considere uma urna com duas bolas brancas e três bolas vermelhas. Dessa urna, são extraídas, sem reposição, duas bolas. Seja X: "número de bolas vermelhas obtidas nas duas extrações". Determine a distribuição de probabilidade de X .



- Os casos possíveis de retirar duas bolas da urna sem reposição são: $\Omega = (B, B), (B, V), (V, B), (V, V)$.



- Seja X: "número de bolas vermelhas obtidas nas duas extrações". Logo os valores que a variável aleatória (X) pode assumir são $x = \{0, 1, 2\}$.
- Então a probabilidade de cada x vai ser:

$$P(x) = P(X = x) \qquad \forall x \in \{0, 1, 2\}$$

$$P(0) = P(X = 0) = \frac{1}{\binom{5}{2}} = \frac{1}{10}$$

$$P(1) = P(X = 1) = \frac{\binom{3}{1}\binom{2}{1}}{\binom{5}{2}} = \frac{6}{10}$$

$$P(0) = P(X = 0) = \frac{\binom{3}{2}}{\binom{5}{2}} = \frac{3}{10}$$

- Só para termos certeza que é realmente uma distribuição de probabilidade, vamos ver se ela satisfaz as duas condições de distribuição de probabilidade.

1. Nitidamente percebemos que os valores são $0 \le x_i \le 1$.
2. Vamos somar todos os valores para ver se resulta em 1.

$$\sum_{x=0}^{2} P(x) = \frac{1}{10} + \frac{6}{10} + \frac{3}{10} = 1$$

- Logo a distribuição de probabilidade vai ser

x	0	1	2
P(x)	1/10	6/10	3/10

▼ **Ex 2**

- Número de filhos - Com dados do último censo, a assistente social de um centro de saúde constatou que para as famílias de uma determinada região:

$$\begin{aligned} &\bullet\ 20\% \text{ não têm filhos} \\ &\bullet\ 30\% \text{ têm } 1 \text{ filho} \\ &\bullet\ 35\% \text{ têm } 2 \text{ filhos} \\ &\bullet\ 15\% \text{ têm igualmente } 3,\ 4 \text{ ou } 5 \text{ filhos} \end{aligned}$$

- Vamos definir a variável aleatória $N =$ "número de filhos por família".
- Suponha que uma família será selecionada aleatoriamente nessa região e o número de filhos averiguado, queremos estudar a v.a N .

Função de Distribuição Acumulada (f.d.a)

- O conceito de Função de Distribuição Acumulada (ou **Distribuição Cumulativa de Probabilidade** ou **Função de Distribuição**) que introduziremos aplica-se tanto a variáveis aleatórias discretas quanto a variáveis aleatórias contínuas.
- A função de distribuição acumulada nos dá uma outra maneira de descrever como as probabilidades são associadas aos valores de uma variável aleatória.

▼ **Definição**

- A função de distribuição acumulada (f.d.a) de uma variável aleatória (discreta ou contínua) X é definida como

$$F(x) = \mathbb{P}(X \le x) \qquad \text{para todo } x$$

- **Propriedades:** Uma função de distribuição acumulada tem que satisfazer as seguintes condições:

$(F1)$ $0 \le F(x) \le 1$

$(F2)$ F é **não decrescente** , isto é, $F(x) \le F(y)$ sempre que $x \le y, \ \forall x, y \in \mathbb{R}$.

$(F3)$ Continuidade à direita. Se $x_n \downarrow x$, então $F(x) \downarrow F(y)$.

$(F4)$ $F(-\infty) = \lim_{x \to -\infty} F(x) = 0 \quad e \quad F(\infty) = \lim_{x \to \infty} F(x) = 1$.

▼ **Demostração das Propriedades**

- $F1$: Como $F(x)$ representa uma probabilidade segue-se que $0 \le F(x) \le 1$.
- $F2$: Note que $[X \le x] \cap [X \le y]$ sempre que $x \le y$. Logo as probabilidades satisfazem à desigualdade:

$$F(x) = P(X \le x) \le P(Xy) = F(y).$$

- $F3$: Seja $x \in \mathbb{R}$ e considere uma sequência $\{x_n, n \ge 1, n \in \mathbb{N}\}$ tal que $x_n \downarrow x$. Isto é, os x_n 's se aproximam de x pela direita ou, em outras palavras, por valores superiores a x. Então, $[X \le x_n] \downarrow [X \le x]$ e, assim, $P(X \le x_n) \downarrow P(X \le x)$. Como o resultado vale para qualquer x, a propriedade está verificada.

- $F4$: Aplicamos a continuidade da probabilidade. Observe que para $X_n \downarrow -\infty$, os eventos$[X \leq x_n] = \{w \in \Omega : X(w) \leq x_n\}$ têm como limite o conjunto vazio. Logo $F(x_n) = P(X \leq x_n) \downarrow 0$. De modo análogo, tomando $X_n \uparrow \infty$, os eventos $[X \leq x_n] \uparrow \Omega$ e, portanto $F(x_n) = P(X \leq x_n) \uparrow 1$.

▼ De Distribuição de Probabilidade para f.d.a

- Quando quisemos encontrar a **f.d.a** de uma **v.a.**, e o problema der a Distribuição de Probabilidade, podemos encontrar a **f.d.a** a partir da Distribuição de Probabilidade por meio da expressão:

$$F(x) = \sum_{i:x_i \leq x} P[X = x_i]$$

- O inverso também é possível:

$$P[X = x_i] = F(x_i) - \lim_{x \to -x_i} F(x)$$

▼ Exemplos

▼ Ex 1:

- Um grupo de 1000 crianças foi analisado para determinar a efetividade de uma vacina contra um tipo de alergia. As crianças recebiam uma dose de vacina e após um mês passavam por um novo teste. Caso ainda tivessem tido alguma reação alérgica, recebiam outra dose. Ao fim de 5 doses, foram consideradas imunizadas.
- Variável de interesse: $X = \text{"número de doses"}$

Doses (X)	1	2	3	4	5
Freq.	245	288	256	145	66

- Uma criança é sorteada ao acaso, qual será a probabilidade dela ter recebido 2 doses?

$$P(X = 2) = \frac{288}{1000} = 0,288$$

Doses (X)	1	2	3	4	5
$p(x)$	0,245	0,288	0,256	0,145	0,066

- Qual a probabilidade da criança ter recebido até duas doses?

$$F(2) = P(X \leq 2) = P(X = 1) + P(X = 2) = 0,245 + 0,288 = 0,533$$

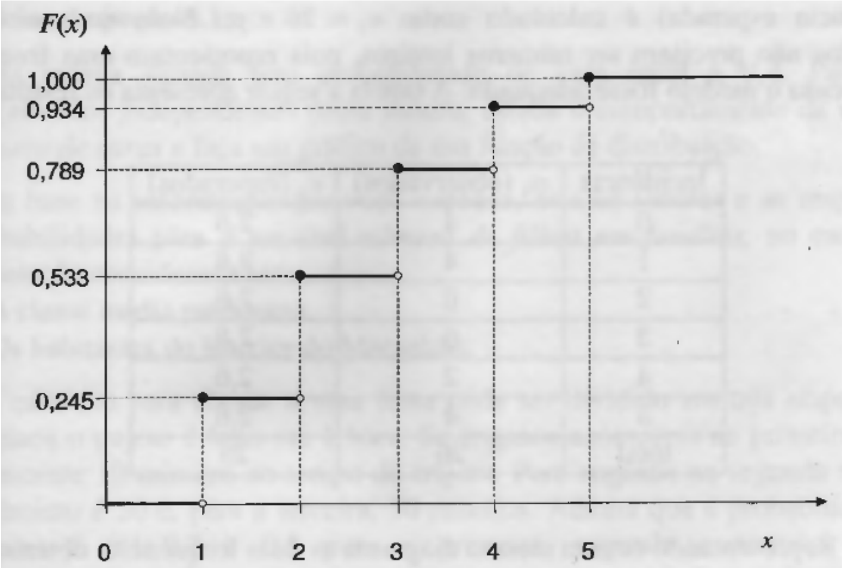
- Note que, tendo em vista que a variável só assume valores inteiros, esse valor fica inalterado no intervalo $[2, 3)$. Isto é , $F(2) = F(2,1) = F(2,45) = F(2,99)$.
- Por essa razão escrevemos:

$$F(x) = P(X \leq x) = 0,533 \quad para \quad 2 \leq x < 3$$

- Logo os valores completos da função de distribuição são os seguintes:

$$F(x) = \begin{cases} 0 & \text{se } x < 1; \\ 0{,}245 & \text{se } 1 \leq x < 2; \\ 0{,}533 & \text{se } 2 \leq x < 3; \\ 0{,}789 & \text{se } 3 \leq x < 4; \\ 0{,}934 & \text{se } 4 \leq x < 5; \\ 1 & \text{se } x \geq 5. \end{cases}$$

- A figura a seguir apresenta um diagrama dessa função:



Esperança de v.a. Discretas ($\mathbb{E}(X)$)

▼ Definição

- **Definição:** Seja X uma variável aleatória discreta, com função de probabilidade $p(x) = P(X = x)$, $x = 1, 2, ..., n, ...$. Então, a esperança (ou **valor esperado** ou **média**) de X, denotada por $\mathbb{E}(X)$ é definida como:

$$\mathbb{E}(X) = \mu x = \sum_{x=1}^{n} x \, \mathbb{P}(X = x) = \sum_{x=1}^{n} x \, p(x) \qquad \forall \, x = 1, 2, 3, \cdots, n$$

- A Esperança matemática também pode ser chamada de **valor médio**.
- Podemos usar o simbolo E para esperança quando houver mais de uma variável envolvida. Ou podemos usar o simbolo grego **mi** μ, quando houver só uma variável.
- Colocando em palavras, o valor esperado de X é uma **média ponderada** dos possíveis valores que X pode receber, com cada valor sendo ponderado pela probabilidade de que X seja igual a esse valor. Por exemplo, se a função de probabilidade de X é dada por:

$$p(0) = \frac{1}{2} = p(1)$$

então

$$E[X] = 0\left(\frac{1}{2}\right) + 1\left(\frac{1}{2}\right) = \frac{1}{2}$$

Nesse caso acima é somente a média ordinária dos dois valores possíveis, 0 e 1, que X pode assumir. Por outro lado, se

$$p(0) = \frac{1}{3} \qquad p(1) = \frac{2}{3}$$

então

$$E[X] = 0\left(\frac{1}{3}\right) + 1\left(\frac{2}{3}\right) = \frac{2}{3}$$

- Agora suponha que conheçamos uma variável aleatória discreta e sua função de probabilidade e que queiramos calcular o valor esperado de alguma função de X ,
digamos, $g(X)$. Como podemos fazer isso?
- Uma maneira é a seguinte: como $g(X)$ é ela mesmo uma variável aleatória discreta, ela tem uma função de probabilidade, que pode ser determinada a partir da função de probabilidade de X. Uma vez que tenhamos determinado a função de

probabilidade de$g(X)$, podemos calcular $E[g(x)]$ usando a definição de valor esperado.

- **Exemplo:** Seja X uma variável aleatória que pode receber os valores $-1, 0, 1$ com respectivas probabilidades:

$$P\{X = -1\} = 0{,}2 \quad P\{X = 0\} = 0{,}5 \quad P\{X = 1\} = 0{,}3$$

Calcule $E[X^2]$

Solução: Seja $Y = X^2$. Então a função de probabilidade de Y é dada por

$$\begin{aligned} P\{Y = 1\} &= P\{X = -1\} + P\{X = 1\} = 0{,}5 \\ P\{Y = 0\} &= P\{X = 0\} = 0{,}5 \end{aligned}$$

Logo

$$E[X^2] = E[Y] = 1(0{,}5) + 0(0{,}5) = 0{,}5$$

Observe que

$$0{,}5 = E[X^2] \neq (E[X])^2 = 0{,}01$$

> Embora o procedimento anterior sempre nos permita calcular o valor esperado de qualquer função de X a partir do conhecimento da função de probabilidade de X, há uma outra maneira de raciocinar sobre $E[g(X)]$: já que $g(X)$ será igual a $g(x)$ sempre que X for igual a x , parece razoável que $E[g(X)]$ deva ser uma média ponderada dos valores $g(x)$, com $g(x)$ sendo ponderado pela probabilidade de que X seja igual a x. Isto é, o resultado a seguir é bastante intuitivo:
>
> $$E[g(X)] = \sum_i g(x_i)p(x_i)$$

- O exemplo acima refeito com esse novo método:

$$\begin{aligned} E\{X^2\} &= (-1)^2(0{,}2) + 0^2(0{,}5) + 1^2(0{,}3) \\ &= 1(0{,}2 + 0{,}3) + 0(0{,}5) \\ &= 0{,}5 \end{aligned}$$

- **Obs:** O uso da denominação **média** para o valor esperado da variável tem origem histórica, mas também pode ser visto como uma referência a um resultado importante conhecido por **Lei dos Grandes Números.**
- **Momento:**
  - **Momento Central de Ordem k:** Para $k = 1, 2, \cdots$, o momento de ordem k da variável X é definido por $E(X^k)$, desde que essa quantidade exista.

    Se $E(X) = \mu < \infty$, definimos o momento central de ordem k por $E[(X - \mu)^k]$, sempre que essa quantidade exista.
  - **Momento Absoluto de Ordem k:** De modo similar, o momento absoluto de ordem k da variável aleatória X é definido por $E(|X|^k)$.
  - A existência dos momentos está relacionada à discussão que fizemos sobre a existência da integral. Isto é, o valor infinito pode ocorrer, mas operações do tipo $\infty - \infty$ são indeterminadas e levam à **inexistência** do momento.
  - É imediato notar que o momento de ordem 1 é o valor esperado, já definido anteriormente.

Propriedades

- Nas seguintes propriedades, X e Y é uma variável aleatória e a, b são constantes.

  $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$

  $\mathbb{E}(X - Y) = \mathbb{E}(X) - \mathbb{E}(Y)$

- $\mathbb{E}(a) = a$
- $\mathbb{E}(a + X) = a + \mathbb{E}(X)$
- $\mathbb{E}(bX) = b\mathbb{E}(X)$
- $\mathbb{E}(a + bX) = a + b\mathbb{E}(X)$
- Para $r \geq 1$, o r-ésimo momento de uma variável aleatória é definido por $\mathbb{E}(X^r)$ (se existe).

Variância de uma v.a. ($Var(X)$)

Definição

- Dada uma variável aleatória X e sua função distribuição F, seria extremamente útil se pudéssemos resumir as propriedades essenciais de F em certas medidas convenientemente definidas.
- Uma dessas medidas seria $E[X]$ o , valor esperado de X.
- Entretanto, embora $E[A]$ forneça a média ponderada dos valores possíveis de X, ela não nos diz nada sobre a variação, ou dispersão, desses valores.
- **Por exemplo**, embora as v.a W, Y e Z com funções discretas de probabilidade determinadas por

$$\begin{aligned} W &= 0 \quad \text{com probabilidade } 1 \\ Y &= \begin{cases} -1 & \text{com probabilidade } \frac{1}{2} \\ +1 & \text{com probabilidade } \frac{1}{2} \end{cases} \\ Z &= \begin{cases} -100 & \text{com probabilidade } \frac{1}{2} \\ +100 & \text{com probabilidade } \frac{1}{2} \end{cases} \end{aligned}$$

  tenham todas a mesma esperança, que é igual a 0. Existe uma dispersão muito maior nos valores possíveis de Y do que naqueles de W (que é uma constante) e nos valores possíveis de Z do que naqueles de Y.
- Como esperamos que X assuma valores em torno de sua média $E[X]$, parece razoável que uma maneira de medir a possível variação de X seja ver, em média, quão distante X estaria de sua média.
- Uma possível maneira de se medir essa variação seria considerar a grandeza $E[|X - \mu|]$, onde $\mu = E[X]$. Entretanto, a manipulação dessa grandeza seria matematicamente inconveniente. Por esse motivo, uma grandeza mais tratável é usualmente considerada $E[(X - \mu)^2]$ esta é a esperança do quadrado da diferença entre X e sua média. Temos assim a definição a seguir.
- **Definição:** Se X é uma variável aleatória com média μ, então a variância de X, representada por $Var(X)$, é definida como:

$$Var(X) = \sigma^2 = E\left[(X - \mu)^2\right] = E\left[\left(X - E(X)\right)^2\right] = E(X^2) - E(X)^2$$

- Denominamos **Desvio-Padrão** de X a raiz quadrada positiva de $Var(X)$, e é denotada por $DP(X)$.

$$DP(X) = +\sqrt{Var(X)}$$

Propriedades

- Nas seguintes propriedades, X é uma variável aleatória e a, b são constantes.

  $Var(X) = E[(X - \mu)^2]$

- $\mathrm{Var}(a) = 0$
- $\mathrm{Var}(a + X) = \mathrm{Var}(X)$
- $\mathrm{Var}(bX) = b^2\mathrm{Var}(X)$
- $\mathrm{Var}(a + bX) = b^2\mathrm{Var}(X)$
- $\mathrm{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$

- **Demostração** da quarta propriedade $Var(a + bX) = b^2 Var(X)$

  Sendo $E(X) = \mu$, por uma propriedade do valor esperado temos $E(aX + b) = a + b\ E(X) = a + b\mu$. Assim,

$$Var(a + bX) = E\left[\Big((a + bX) - (a + b\mu)\Big)^2\right] = E\left[\Big(a + bX - a - b\mu\Big)^2\right] = E\left[\Big(bX - b\mu\Big)^2\right] = b^2 E\left[\Big(X - \mu\Big)^2\right] = b^2 Var(X)$$

- **Demostração** da última propriedade $Var(X) = E(X^2) - \big[E(X)\big]^2$.

  Para verificar, vamos desenvolver o quadrado do desvio e aplicar propriedades da esperança.

$$Var(X) = E\left[(X - \mu)^2\right] = E\left[X^2 - 2X\mu + \mu^2\right] = E(X^2) - E(2X\mu) + E(\mu^2) = E(X^2) - 2\mu E(X) + \mu^2 = E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2 = E(X^2) - E(X)^2 = Var(X)$$

Resumo das distribuições discretas

Distribuição	$P(X = x)$	Parâmetros	$E(X)$	$Var(X)$	f.g.p	f.g.m
Uniforme Discreta	$\frac{1}{N},\quad x = 1, 2, \cdots, N$	N	$\frac{N+1}{2}$	$\frac{N^2-1}{2}$	–	–
Bernoulli	$p^x(1-p)^{1-x},\quad x = 0, 1$	p	p	$p(1-p)$	–	–
Binomial	$\binom{n}{x}p^x(1-p)^{n-x},\quad x = 0, 1, ..., n$	n, p	np	$np(1-p)$	$[pt + (1-p)]^n$	$(pe^t + 1 - p)^n$
Geométrica	$p(1-p)^{x-1},\quad x = 1, 2, 3, \cdots$	p	$\frac{1}{p}$	$\frac{1-p}{p^2}$	$\frac{pt}{1-t(1-p)}$	$\frac{pe^t}{1-e^t(1-p)}$
Binomial Negativa	$\binom{x-1}{r-1}p^r(1-p)^{x-r},\quad x = r, (r+1), (r+2), \cdots$	r, p	$\frac{r}{p}$	$\frac{r(1-p)}{p^2}$	–	$\left[\frac{pe^t}{1-e^t(1-p)}\right]^r$
Hipergeométrica	$\frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}},\quad x = 0, 1, 2, \cdots$	N, n, r	$\frac{nr}{N}$	$\frac{nr}{N}(1 - \frac{r}{N})(\frac{N-n}{N-1})$	–	–
Poisson	$\frac{e^{-\lambda}\lambda^x}{x!},\quad x = 0, 1, 2, \cdots$	λ	λ	λ	$e^{\lambda(t-1)}$	$e^{\lambda(e^t-1)}$

Distribuição Uniforme Discreta

▼ **Definição**

- Este é o caso mais simples de v.a. discreta, em que cada valor possível ocorre com a mesma probabilidade.
- Considere uma v.a X que toma um número finito de valores com a mesma probabilidade, ou seja, $X \in \{1, 2, ..., N\}$.
- Dessa forma podemos notar que,

$$p(x) = P(X = x) = \frac{1}{N} \quad \text{para todo} \quad x \in \{1, 2, ..., N\}$$

- Uma v.a com essas propriedades é dita ter uma distribuição **uniforme discreta** com parâmetro N.
- **Notação:** $X \sim U\{1, 2, ..., N\}$.
- Pode se mostrar que:

$$\mathbb{E}(X) = \frac{N+1}{2}$$

$$Var(X) = \frac{N^2-1}{2}$$

▼ **Exemplo**

▼ **Ex 1:** Lançamos um dado equilibrado e observamos a face que ocorreu. A v.a sera $X =$ "resultado obtido no lançamento de um dado honesto".

- Como todos os resultados de X são equiprováveis:

x	1	2	3	4	5	6
$p(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

- Podemos então usar a distribuição uniforme discreta, que sera: $X \sim U\{1,2,3,4,5,6\}$.
- A **esperança** desse experimento vai ser:

$$E(X) = \frac{1}{6} \cdot (1+2+3+4+5+6) = \frac{21}{6} = 3,5$$

Poderíamos ter usado a formula de esperança para distribuição uniforme discreta:

$$E(X) = \frac{N+1}{2} = \frac{7}{2} = 3,5$$

- A **variância** desse experimento vai ser:

$$Var(X) = \frac{1}{6} \cdot \left[(1+4+9+16+25+36) - \frac{1}{6} \cdot (21)^2\right] = \frac{35}{2} = 17,5$$

Poderíamos ter usado a formula de variância para distribuição uniforme discreta:

$$E(X) = \frac{N^2-1}{2} = \frac{6^2-1}{12} = \frac{35}{2} = 3,5$$

Modelo Bernoulli

▼ **Definição**

- Em muitas aplicações, cada observação de um experimento aleatório é binária: tem apenas dois resultados possíveis.
- Por exemplo, uma pessoa pode:

  - aceitar ou recusar uma oferta de cartão de crédito de seu banco.
  - ter ou não um plano de saúde.
  - votar sim ou não em uma assembléia.

- Nesse caso, os resultados do experimento podem ser classificados como sucesso ou fracasso. Exemplos:

  - Lançar uma moeda e verificar a ocorrência de cara ou coroa. Podemos considerar como sucesso, a obtenção de cara.
  - Lançar um dado e verificar se o resultado é par ou ímpar. Podemos considerar como sucesso obter um número par.

- Esse tipo de experimento é denominado **ensaio de Bernoulli**. Ou seja se o experimento tem resposta dicotômica do tipo sucesso-fracasso, ele vai ser denominado ensaio de bernoulli.

Distribuição Bernoulli

Definição

- Experimentos que resultam numa v.a. de Bernoulli são chamados **ensaios de Bernoulli**.
- Para indicar a distribuição de Bernoulli com parâmetro p de uma v.a., usaremos a notação:

$$X \sim Ber(p)$$

O p é a probabilidade de ocorrer um evento o outro.

- Denotaremos com p a probabilidade de sucesso e 1-p a probabilidade de fracasso:

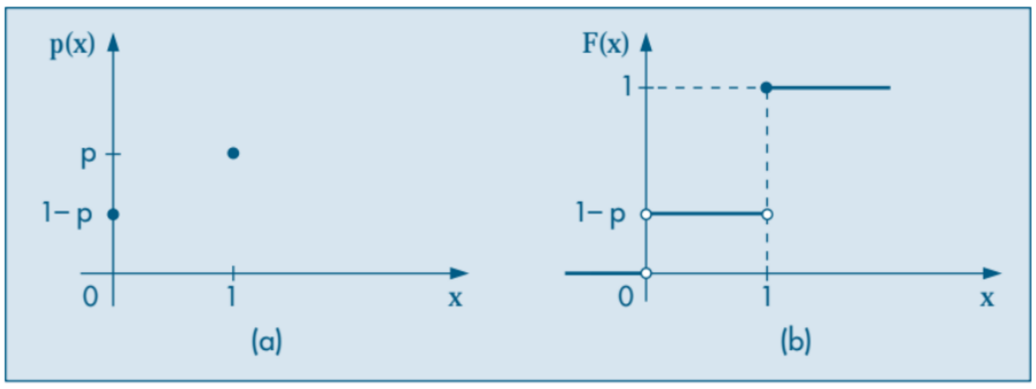
$$P(X=1) = P(sucesso) = p$$
$$P(X=0) = P(fracasso) = 1-p$$

- Dizemos que X tem distribuição de Bernoulli com parâmetro p se sua função de probabilidade é dada por:

$$P(X=x) = p^x(1-p)^{1-x} \qquad \text{em que} \qquad x = 0 \; ou \; 1$$

- Logo por definição temos as formulas para esperança, variância e f.d.a:

  - $\mathbb{E}(X) = 0 \times (1-p) + 1 \times p = p$
  - $\mathbb{E}(X^2) = 0^2 \times (1-p) + 1^2 \times p = p$
  - $\mathrm{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = p - p^2 = p(1-p)$
  - $\text{f.d.a.: F(x)} = \begin{cases} 0, & x < 0 \\ 1-p, & x \in [0,1) \\ 1, & x \geq 1 \end{cases}$



Exemplo

- **Ex 1:** Lançamos um dado e consideramos como sucesso, a obtenção da face 5. Supondo que o dado é honesto:

  Logo o p vai ser igual a probabilidade de sair face 5: $p = \frac{1}{6}$

x	0	1
$p(x)$	$\frac{5}{6}$	$\frac{1}{6}$

Vamos prova que a probabilidade se não sair a face 5:

$$P(X=0) = p^0(1-p)^{1-0} = \left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^{1-0} = \frac{5}{6}$$
$$P(X=1) = p^1(1-p)^{1-1} = \left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^{1-1} = \frac{1}{6}$$

Vamos calcular a $E(X)$ e a $Var(X)$:

Pela definição $E(X) = p$ e $Var(X) = p(1-p)$, então temos:

$$E(X) = \frac{1}{6}$$
$$Var(X) = \frac{1}{6}\left(1-\frac{1}{6}\right) = \frac{1}{6}\cdot\frac{5}{6} = \frac{5}{36}$$

Distribuição Binomial

▼ Definição

- A repetição de um ensaios de Bernoulli qualquer, onde cada repetição é **independente** da outra, dar origem a uma **variável aleatória Binomial**.
  - **Definição:** Suponha que n ensaios de Bernoulli independentes, são realizados. Ou seja, em cada ensaio pode acontecer sucesso com probabilidade p ou fracasso com probabilidade $1-p$.
- Defina a v.a X como o **número de sucessos em n ensaios**, então X é dita ter distribuição binomial com parâmetros n e p e sua função de probabilidade é dada por:

$$p(x) = P(X=x) = \binom{n}{x}p^x(1-p)^{n-x} \qquad \forall \qquad x = \{0, 1, ..., n\}$$

- **Notação:** $X \sim B(n, p)$.
- Note que, $Ber(p) \equiv B(1, p)$.
- $f.d.a\ P(X \leq x) = \sum_{i=0}^{\lfloor x \rfloor}\binom{n}{i}p^i(1-p)^{n-i}$
- Se $X \sim B(n, p)$ então:
  - $E(X) = np$
  - $Var(X) = np(1-p)$
  - $E(X^k) = np\, E[(Y+1)^{k-1}]$ sendo que $Y \sim B(n-1, p)$.
- **Demostração da $E(X)$, $E(X^2)E(X^k)$, $Var(X)$.**
  - ▼ $E(X)$:

$$E(X) = \sum_{x=0}^{n} x\cdot p(x) = \sum_{x=1}^{n} x\cdot\binom{n}{x}p^x(1-p)^{n-x}$$

Note que, quando $x = 0$ a equação vai ser igual a 0, logo o somatório começa no 1.

Note também:

$$x\cdot\binom{n}{x} = x\cdot\frac{n!}{x!(n-x)!} = x\cdot\frac{n(n-1)!}{x(x-1)!(n-x)!} = \frac{n(n-1)!}{(x-1)!(n-x)!} = n\binom{n-1}{x-1}$$

Logo:

$$\sum_{x=1}^{n} x\cdot\binom{n}{x}p^x(1-p)^{n-x} = \sum_{x=1}^{n} n\binom{n-1}{x-1}p^{x-1}\ p\ (1-p)^{(n-1)-(x-1)}$$
$$y = x-1 \qquad m = n-1$$
$$np\sum_{y=0}^{m}\binom{m}{y}p^y\ (1-p)^{m-y} = np\cdot 1 = np$$

Note que essa última somatória da equação acima é a própria função de probabilidade da binomial que é igual a 1.

▼ $E(X^2)$:

Sabemos que $E(X) = np$ e para $E(X^2)$ vem que

$$\begin{aligned}
\mathbb{E}(X^2) &= \sum_{x=0}^{n} x^2 \binom{n}{x} p^x (1-p)^{n-x} \\
&\text{*Note que } x^2 = [x(x-1)+x] \\
&= \sum_{x=0}^{n} [x(x-1)+x] \binom{n}{x} p^x (1-p)^{n-x} \\
&= \left(\sum_{x=0}^{n} x(x-1) \binom{n}{x} p^x (1-p)^{n-x} \right) + \left(\sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} \right) \\
&= \left(\sum_{x=0}^{n} x(x-1) \binom{n}{x} p^x (1-p)^{n-x} \right) + \mathbb{E}(X) \\
&\text{*Note que, o somatorio deve começar no 2} \\
&= \mathbb{E}(X) + \sum_{x=2}^{n} x(x-1)\ \frac{n!}{x!(n-x)!}\ p^x (1-p)^{n-x} \\
&= \mathbb{E}(X) + \sum_{x=2}^{n} x(x-1)\ \frac{n(n-1)(n-2)!}{x(x-1)(x-2)!(n-x)!}\ p^x (1-p)^{n-x} \\
&= \mathbb{E}(X) + n(n-1) \sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!(n-x)!}\ p^{x-2} \cdot p^2 (1-p)^{(n-2)-(x-2)} \\
&= \mathbb{E}(X) + n(n-1)p^2 \sum_{x=2}^{n} \binom{n-2}{x-2} p^{x-2} (1-p)^{(n-2)-(x-2)} \\
&\text{*}j = x-2 \quad \forall\, j = 0,1,2,3,\cdots,n-2 \\
&= \mathbb{E}(X) + n(n-1)p^2 \sum_{j=0}^{n-2} \binom{n-2}{j} p^{j} (1-p)^{(n-2)-j} \\
&\text{*Note que o somatório é uma função de probabilidade logo ele é igual a 1} \\
&= \mathbb{E}(X) + n(n-1)p^2 \cdot 1 \\
&= np + (n^2 - n)p^2 \\
&= np + n^2p^2 - np^2 \\
&= \mathbf{n^2p^2 + np(1-p)}
\end{aligned}$$

▼ $E(X^k)$:

$$E(X^k) = \sum_{x=0}^{n} x^k \cdot p(x) = \sum_{x=1}^{n} x^k \cdot \binom{n}{x} p^x (1-p)^{n-x}$$

Note que, quando $x = 0$ a equação vai ser igual a 0, logo o somatório começa no 1.

Note também:

$$x^k \cdot \binom{n}{x} = x^k \cdot \frac{n!}{x!(n-x)!} = x^k \cdot \frac{n(n-1)!}{x(x-1)!(n-x)!} = x^{k-1} \frac{n(n-1)!}{(x-1)!(n-x)!} = x^{k-1}\, n \binom{n-1}{x-1}$$

Logo:

$$\sum_{x=1}^{n} x^k \cdot \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=1}^{n} x^{k-1}\, n \binom{n-1}{x-1} p^{x-1}\, p\, (1-p)^{(n-1)-(x-1)} =$$
$$np \sum_{x=1}^{n} x^{k-1} \binom{n-1}{x-1} p^{x-1}\, (1-p)^{(n-1)-(x-1)}$$
$$y = x-1 \qquad m = n-1 \qquad x = y+1 \qquad \forall\, y = 0,1,2,3,\cdots,m$$
$$np \sum_{y=0}^{m} (y+1)^{k-1} \binom{m}{y} p^y\, (1-p)^{m-y} = np \sum_{y=0}^{m} (y+1)^{k-1}\, p(x)$$
$$= np \cdot E\Big[(Y+1)^{k-1}\Big]$$

Concluímos então:

$$E(X^k) = np \cdot E\Big[(Y+1)^{k-1}\Big] \qquad\qquad Y \sim B(n-1;p)$$

Usando essa formula para calcular $E(X^2)$:

$$E(X^2) = np \cdot E(Y+1) = np \cdot [(n-1)p + 1]$$

▼ $Var(X)$

$$\begin{aligned} Var(X) &= \mathbb{E}(X^2) - \mu^2 \\ &= n^2p^2 + np(1-p) - (np)^2 \\ &= \mathbf{np(1-p)} \end{aligned}$$

- **Proposição:**

Se $X \sim B(n,p)$ então

$$\mathbb{P}(X = x+1) = \frac{p}{1-p}\frac{n-x}{x+1}\mathbb{P}(X = x)$$

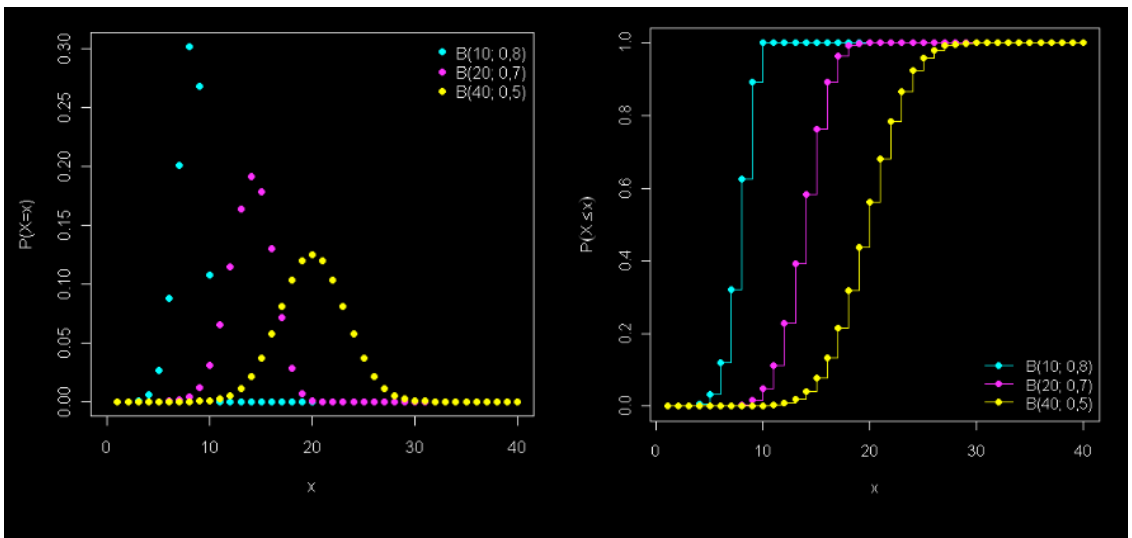


Figura: Representação gráfica da distribuição binomial

▼ **Exemplo**

▼ **Ex 1:** Sabe-se que a eficiência de uma vacina é de **80%**. Um grupo de 3 indivíduos é sorteado, dentre a **população vacinada**, e cada um é submetido a testes para averiguar se está imunizado. Nesse caso, consideramos como sucesso a imunização.

$$X_i = \begin{cases} 1, & \text{indivíduo } i \text{ está imunizado} \\ 0, & \text{caso contrário} \end{cases}$$

- Pelo enunciado, sabe-se que $P(X_i = 1) = p = 0,8$.
- Os três $\{1,2,3\}$ indivíduos são independentes, cada um vai ter uma v.a $\{X_1, X_2, X_3\}$, que assumem valor 1 se o indivíduo está imunizado ou 0 caso contrário.

- A v.a X desse experimento vai ser igual ao número de indivíduos imunizados no grupo, X poderá assumir valores $\{0, 1, 2, 3\}$.
- Note que $X = X_1 + X_2 + X_3$.
- Logo o número de casos possíveis são:

evento	$\mathbb{P}$(evento)	X
$X_1 = 0,\ X_2 = 0,\ X_3 = 0$	$(0,2)^3$	0
$X_1 = 1,\ X_2 = 0,\ X_3 = 0$	$0,8 \times (0,2)^2$	1
$X_1 = 0,\ X_2 = 1,\ X_3 = 0$	$0,8 \times (0,2)^2$	1
$X_1 = 0,\ X_2 = 0,\ X_3 = 1$	$0,8 \times (0,2)^2$	1
$X_1 = 1,\ X_2 = 1,\ X_3 = 0$	$(0,8)^2 \times 0,2$	2
$X_1 = 1,\ X_2 = 0,\ X_3 = 1$	$(0,8)^2 \times 0,2$	2
$X_1 = 0,\ X_2 = 1,\ X_3 = 1$	$(0,8)^2 \times 0,2$	2
$X_1 = 1,\ X_2 = 1,\ X_3 = 1$	$(0,8)^3$	3

- Assim, as probabilidades de cada valor possível de X são:

x	0	1	2	3
$p(x)$	$(0,2)^3$	$3 \times 0,8 \times (0,2)^2$	$3 \times (0,8)^2 \times 0,2$	$(0,8)^3$

- E o comportamento de X é completamente determinado pela função:

$$p(x) = P(X = x) = \binom{3}{x}(0,8)^x(1-0,8)^{n-x} \qquad \forall \qquad x = \{0, 1, 2, 3\}$$

Distribuição Hipergeométrica

▼ **Definição**

$$p(x) = \mathbb{P}(X = x) = \frac{\begin{pmatrix} r \\ x \end{pmatrix}\begin{pmatrix} N-r \\ n-x \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}}$$

- Este é um modelo para amostragem sem reposição de uma população com um número finito de elementos que podem ser classificados em duas categorias mutuamente excludentes.
- **Detalhes:**
    - N objetos
    - r possuem uma característica A
    - $N - r$ possuem uma característica B
    - Um grupo de n elementos é escolhido ao acaso, dentre os N possíveis, sem reposição.
- **Objetivo:** Calcular a probabilidade de que este grupo de n elementos contenha x elementos com a característica A.
- Seja X a v.a representando o número de elementos com a característica A, dentre os $n$elementos selecionados. Dizemos que X tem distribuição hipergeométrica e sua distribuição de probabilidade é dada por:

$$p(x) = \mathbb{P}(X = x) = \frac{\begin{pmatrix} r \\ x \end{pmatrix} \begin{pmatrix} N - r \\ n - x \end{pmatrix}}{\begin{pmatrix} N \\ n \end{pmatrix}} \quad \forall$$

$$\max\{0, n - (N - r)\} \leq x \leq \min\{r, n\}.$$

- **Notação:** $X \sim Hip(N, n, r)$
- Se $X \sim Hip(N, n, r)$ então:
  - $\mathbb{E}(X) = n\frac{r}{N}$
  - $Var(X) = n\frac{r}{N}(1 - \frac{r}{N})(\frac{N-n}{N-1})$

Distribuição Geométrica

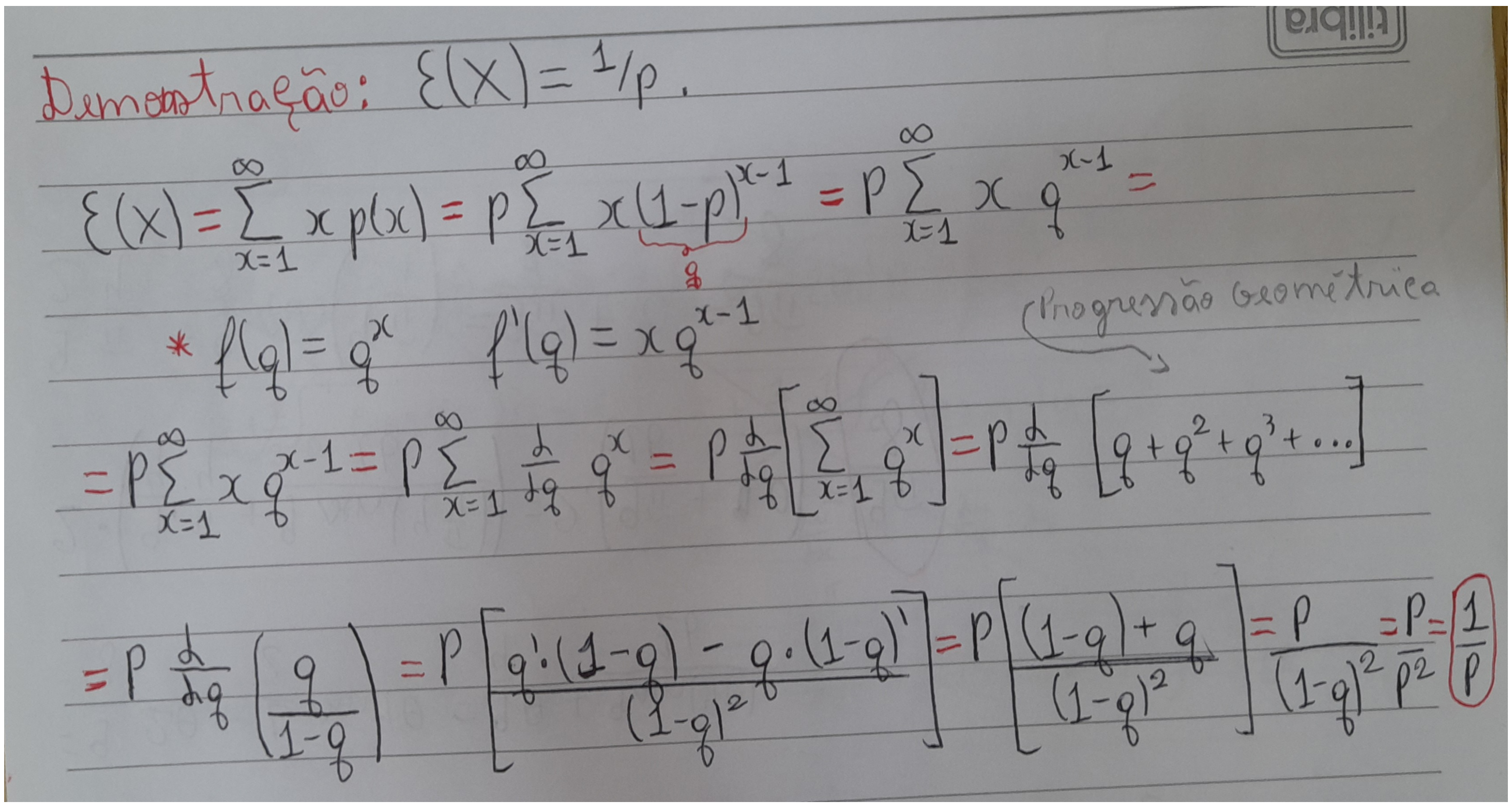
▼ Definição

- Consideremos uma sequência **ilimitada** de ensaios de Bernoulli, com probabilidade de sucesso p em cada ensaio.
- Realizamos os ensaios até que ocorra o primeiro sucesso.
- Seja X a variável aleatória que conta o número de ensaios até o primeiro sucesso.
- Então X segue Distribuição Geométrica com parâmetro $0 < p < 1$ e a sua função de probabilidade é dada por:

$$\mathbb{P}(X = x) = p(1 - p)^{x-1} \text{ para todo } x = 1, 2, 3, \ldots$$

$$F(x) = 1 - (1 - p)^x$$

- **Notação:** $X \sim Geo(p)$.
- **Propriedades:**
  - $E(X) = \frac{1}{p}$
  - $E(X^2) = \frac{2-p}{p^2}$
  - $Var(X) = \frac{1-p}{p^2}$
- **Demostração:**
  - $E(X) = \frac{1}{p}$.



Demonstração: $E(X) = 1/p$.

$$E(X) = \sum_{x=1}^{\infty} x\,p(x) = p\sum_{x=1}^{\infty} x\,\underbrace{(1-p)}_{q}{}^{x-1} = p\sum_{x=1}^{\infty} x\, q^{x-1} =$$

(Progressão Geométrica

$$* \; f(q) = q^{x} \qquad f'(q) = x\,q^{x-1}$$

$$= p\sum_{x=1}^{\infty} x\,q^{x-1} = p\sum_{x=1}^{\infty} \frac{d}{dq} q^{x} = p\frac{d}{dq}\left[\sum_{x=1}^{\infty} q^{x}\right] = p\frac{d}{dq}\left[q+q^2+q^3+\dots\right]$$

$$= p\frac{d}{dq}\left(\frac{q}{1-q}\right) = p\left[\frac{q'\cdot(1-q) - q\cdot(1-q)'}{(1-q)^2}\right] = p\left[\frac{(1-q)+q}{(1-q)^2}\right] = \frac{p}{(1-q)^2} = \frac{p}{p^2} = \boxed{\frac{1}{p}}$$

- $E(X^2) = \frac{2-p}{p^2}$
  - O r-ésimo momento fatorial de uma v.a. X é definido por:

$$\mathbb{E}\Big[(X)_r\Big] = \mathbb{E}\Big[X(X-1)(X-2)...(X-r+1)\Big]$$

  - Por exemplo, para $r = 2$, temos:

$$\mathbb{E}\Big[(X)_2\Big] = \mathbb{E}\Big[X(X-1)\Big] = \sum_{x=1}^{\infty} x(x-1)\mathbb{P}(x)$$

  - Note que:

$$\mathbb{E}\Big[X(X-1)\Big] = \mathbb{E}\Big[X^2 - X\Big] = \mathbb{E}(X^2) - \mathbb{E}(X)$$

  - Logo:

$$\mathbb{E}(X^2) = \mathbb{E}\Big[X(X-1)\Big] + \mathbb{E}(X)$$

  - Agora, vamos encontra o valor de $\mathbb{E}\Big[X(X-1)\Big]$.
  - Pela definição de esperança temos que:

$$\mathbb{E}\Big[X(X-1)\Big] = \sum_{x=1}^{\infty} x(x-1)\mathbb{P}(x)$$

  - Pela definição de função de probabilidade da Distribuição Geométrica, temos que $\mathbb{P}(x) = p(1-p)^{x-1}$, onde p é o parâmetro da Distribuição Geométrica.

$$\sum_{x=1}^{\infty} x(x-1)\mathbb{P}(x) = \sum_{x=1}^{\infty} x(x-1)p(1-x)^{x-1} = p\sum_{x=1}^{\infty} x(x-1)(1-x)^{x-1}$$

  - Agora vamos substituir $(1-p)$ por q e vamos multiplicar e dividir por q:

$$p\sum_{x=1}^{\infty} x(x-1)(1-x)^{x-1} = p\sum_{x=1}^{\infty} x(x-1)\,q^{x-1}\,\frac{q}{q} = pq\sum_{x=1}^{\infty} x(x-1)\,q^{x-2}$$

  - Note que a derivada segunda de q^x é igual a:

$$\frac{d^2}{dq^2}(q^x) = \frac{d}{dq}(xq^{x-1}) = x(x-1)q^{x-2}$$

- Logo temos:

$$pq\sum_{x=1}^{\infty} x(x-1)\,q^{x-2} = pq\sum_{x=1}^{\infty}\frac{d^2}{dq^2}q^x = pq\,\frac{d^2}{dq^2}\sum_{x=1}^{\infty}q^x = pq\,\frac{d^2}{dq^2}\left[q^2+q^3+q^4+...\right]$$

- Note que o somatório é uma PG, logo a soma dos termos de uma PG infinita é data pela formula $\frac{a_1}{1-r}$, nesse caso temos então $\left(\frac{q^2}{1-q}\right)$. Essa fórmula só é válida para progressões geométricas decrescentes, com 0 < q < 1:

$$pq\,\frac{d^2}{dq^2}\left[q^2+q^3+q^4+...\right] = pq\,\frac{d^2}{dq^2}\left(\frac{q^2}{1-q}\right)$$

- Resolvendo a derivada, temos:

$$pq\,\frac{d^2}{dq^2}\left(\frac{q^2}{1-q}\right) = pq\,\frac{2}{(1-q)^3} = \frac{2pq}{(1-q)^3}$$

- Substituindo o q pelo valor original $q = 1-p$:

$$\frac{2pq}{(1-q)^3} = \frac{2p(1-p)}{\left(1-(1-p)\right)^3} = \frac{2p(1-p)}{(1-1+p)^3} = \frac{2p(1-p)}{p^3} = \frac{2(1-p)}{p^2}$$

- Pronto agora vamos voltar para a formula do $\mathbb{E}(X^2)$:

$$\mathbb{E}(X^2) = \mathbb{E}\Big[X(X-1)\Big] + \mathbb{E}(X) = \frac{2(1-p)}{p^2} + \frac{1}{p} = \frac{2(1-p)+p}{p^2} = \frac{2-2p+p}{p^2} = \frac{2-p}{p^2}$$

- $Var(X) = \frac{1-p}{p^2}$.

Demonstração:

$$Var(X) = \mathbb{E}(X^2) - \mathbb{E}^2(X) = \frac{2-p}{p^2} - \left(\frac{1}{p}\right)^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \boxed{\frac{1-p}{p^2}}$$

- **Proposição:**

Seja $X \sim \text{Geo}(p)$. Para quaisquer números inteiros positivos s e t, vale

$$\mathbb{P}(X \geq s+t | X \geq s) = \mathbb{P}(X \geq t).$$

- Essa propriedade é conhecida como **Falta de Memória Geométrica**.
- Ela indica a maneira como a variável incorpora a informação anterior.
- Em termos ilustrativos, podemos considerar que a variável “lembra” do presente, mais “esqueceu” do que ocorreu no passado.

Por exemplo, se X representasse a espera em dias para a ocorrência de um certo evento, a probabilidade condicional acima representaria a probabilidade de espera de pelo menos $m+n$ dias, sabendo que o evento não ocorreu antes de m dias. A propriedade da falta de memória estabelece que essa probabilidade é a mesma de esperar pelo menos n dias.

Para verificar a propriedade, observe que

$$P(X \geq m+n \,|\, X \geq m) = \frac{P([X \geq m+n] \cap [X \geq m])}{P(X \geq m)} = \frac{P(X \geq m+n)}{P(X \geq m)}.$$

A função de probabilidade de X é dada por $P(X=x) = p(1-p)^x$. Então,

$$P(X \geq m+n \,|\, X \geq m) = \frac{(1-p)^{m+n}}{(1-p)^m} = (1-p)^n = P(X \geq n).$$

- Pode ainda demonstrar que o modelo Geométrico é o único modelo discreto com a propriedade da falta de memória.
- Uma generalização do modelo Geométrico é o Modelo Binomial Negativo.

Distribuição Binomial Negativa (Pascal)

▼ **Definição**

- Esta distribuição pode ser considerada como uma generalização da distribuição geométrica.
- Considere uma sequência ilimitada de ensaios de Benoulli, e seja X a v.a. que conta o número de ensaios até a ocorrência de r sucessos.
- Nessas condições X é dita ter distribuição **binomial negativa**.
- E sua função de probabilidade é dada por:

$$\mathbb{P}(x) = \mathbb{P}(X=x) = \binom{x-1}{r-1} p^r (1-p)^{x-r} \qquad \forall \; x = r, r+1, r+2, \ldots$$

- **Notação:** $X \sim BN(r,p)$
- Logo temos:
  - $E(X) = \frac{r}{p}$
  - $Var(X) = \frac{r(1-p)}{p^2}$
- **Obs:** A distribuição geométrica equivale a:

$$Geo(p) \equiv BN(1,p)$$

- A distribuição binomial negativa pode, também, ser definida em termos da v.a Y que conta o número de fracassos antes do r-ésimo sucesso. Esta formulação é equivalente a definição anterior, uma vez que, $Y = X - r$.

$$P(y) = P(Y=y) = \binom{r+y-1}{y} p^r (1-p)^y, \quad \forall \, y = 0, 1, \ldots$$

- Podemos ver que:

$$\sum_{x=r}^{\infty} P(X=x) = \sum_{x=r}^{\infty} \binom{x-1}{r-1} p^r (1-p)^{x-r} = 1 \tag{1}$$

Pela própria definição da Função de Probabilidade fica provado que a equação (1) é igual a 1.

▼ **Demonstração:** Mais podemos mostra que a equação (1) é igual a 1:

- Para podemos fazer essa demonstração, vamos precisar primeiro mostrar alguns teoremas.

- **Teorema Binomial:** $(1+x)^r = 1 + rx + \binom{r}{2}x^2 + \binom{r}{3}x^3 + \ ...$
  - Para provamos essa igualdade, vamos considera $f(x) = (1+x)^r$, e vamos usar a serie de Taylor, mais especificamente a serie de Mac-laurim.
  - Então temos:

$$(1+x)^r = f(0) + f'(0)x + f''(0)\frac{x^2}{2!} + f^3(0)\frac{x^3}{3!} + \cdots + f^r(0)\frac{x^r}{r!}$$

$$= 1 + rx + \frac{r(r-1)}{2!}x^2 + \frac{r(r-1)(r-2)}{3!}x^3 + \cdots$$

note que:

$$\binom{n}{p} = \frac{n!}{p!(n-p)!} = \frac{n(n-1)(n-2)\cdots(n-p+1)(n-p)!}{p!(n-p)!} = \frac{n(n-1)(n-2)\cdots(n-p+1)}{p!}$$

então, usando essa afirmação na equação anterior temos:

$$(1+x)^r = 1 + rx + \frac{r(r-1)}{2!}x^2 + \frac{r(r-1)(r-2)}{3!}x^3 + \cdots$$

$$= 1 + rx + \binom{r}{2}x^2 + \binom{r}{3}x^3 + \cdots$$

logo fica provado o teorema binomial.

- Note que essa serie pode ser representado por um somatório:

$$(1+x)^r = 1 + rx + \binom{r}{2}x^2 + \binom{r}{3}x^3 + \ ... = \sum_{k=0}^{\infty}\binom{r}{k}x^k$$

- Agora vamos demostrar que:
- **Demostração:** $(1-x)^{-r} = \sum_{k=0}^{\infty}\binom{r+k-1}{k}x^k$
  - Sabemos que pelo teorema binomial podemos escrever o $(1-x)^{-r}$ como:

$$(1-x)^{-r} = \sum_{k=0}^{\infty}\binom{-r}{k}(-x)^k$$

para facilitar as contas vamos retira o sinal negativo do x:

$$(1-x)^{-r} = \sum_{k=0}^{\infty}\binom{-r}{k}(-x)^k = \sum_{k=0}^{\infty}(-1)^k\binom{-r}{k}x^k \tag{2}$$

  - Agora vamos **demostrar** que $\binom{-r}{k}(-1)^k = \binom{r+k-1}{k}$
  - Temos que:

$$\binom{-r}{k}(-1)^k = \frac{-r(-r-1)(-r-2)\cdots(-r-k+1)(-r-k)!}{k!(-r-k)!}(-1)^k$$

$$= \frac{-r(-r-1)(-r-2)\cdots(-r-k+1)}{k!}(-1)^k$$

podemos colocar o menos em evidência:

$$\binom{-r}{k}(-1)^k = \frac{-r(-r-1)(-r-2)\cdots(-r-k+1)}{k!}(-1)^k$$

$$= (-1)^k\frac{r(r+1)(r+2)\cdots(r+k-1)}{k!}(-1)^k$$

note que $(-1)^k \cdot (-1)^k = 1$ para todo k impar ou par.

$$\binom{-r}{k}(-1)^k = \frac{r(r+1)(r+2)\cdots(r+k-1)}{k!} \cdot 1$$

agora vamos reescrever a equação acima de traz para frente:

$$\binom{-r}{k}(-1)^k = \frac{r(r+1)(r+2)\cdots(r+k-1)}{k!} \cdot 1$$

$$= \frac{(r+k-1)(r+k-2)\cdots(r+2)(r+1)r}{k!}$$

veja que o termo depois do r é o $(r-1)!$ que foi cancelado com o fatorial do denominador:

$$\binom{-r}{k}(-1)^k = \frac{(r+k-1)(r+k-2)\cdots(r+2)(r+1)r}{k!}$$

$$= \frac{(r+k-1)(r+k-2)\cdots(r+2)(r+1)r\ (r-1)!}{k!(r-1)!}$$

$$= \frac{(r+k-1)!}{k!(r-1)!}$$

$$= \binom{r+k-1}{k}$$

- Agora podemos volta pra a equação (2), relembrando ela era:

$$(1-x)^{-r} = \sum_{k=0}^{\infty}(-1)^k\binom{-r}{k}x^k$$

temos então a equação (3):

$$(1-x)^{-r} = \sum_{k=0}^{\infty}(-1)^k\binom{-r}{k}x^k = \sum_{k=0}^{\infty}\binom{r+k-1}{k}x^k \tag{3}$$

- Pronto, agora podemos prova a equação (1), relembrando ela era:

$$\sum_{x=r}^{\infty}\binom{x-1}{r-1}p^r(1-p)^{x-r} = 1$$

- Se $X \sim BN(r,p)$, então

$$\mathbb{P}(X=x) = \binom{x-1}{r-1}p^r(1-p)^{x-r} \qquad \forall\ x = r, r+1, r+2, \ldots$$

- Então:

$$\sum_{x=r}^{\infty}\binom{x-1}{r-1}p^r(1-p)^{x-r} = p^r\sum_{x=r}^{\infty}\binom{x-1}{r-1}(1-p)^{x-r}$$

$$= p^r\sum_{x=r}^{\infty}\binom{x-1}{r-1}\,q^{x-r} \qquad q=(1-p)$$

- Vamos substituir o $x-r$ por y, e por consequência o $x=y+r$:

$$p^r\sum_{x=r}^{\infty}\binom{x-1}{r-1}\,q^{x-r} = p^r\sum_{x=r}^{\infty}\binom{r+y-1}{r-1}\,q^y$$

- Essa combinação pode ser rescrita como:

$$\binom{r+y-1}{r-1} = \frac{(r+y-1)!}{(r-1)!\,\big[r+y-1-(r-1)\big]!} = \frac{(r+y-1)!}{(r-1)!\,y!} = \binom{r+y-1}{y}$$

- Logo temos:

$$p^r\sum_{x=r}^{\infty}\binom{r+y-1}{r-1}\,q^y = p^r\sum_{x=r}^{\infty}\binom{r+y-1}{y}\,q^y$$

- Pela equação (3) sabemos que:

$$p^r\sum_{x=r}^{\infty}\binom{r+y-1}{y}\,q^y = p^r(1-q)^{-r} = \frac{p^r}{(1-q)^r} = \frac{p^r}{\big(1-(1-p)\big)^r} = \frac{p^r}{p^r} = 1$$

- **Curiosidade:**
  - Pascal ou binomial negativa?
  - Jain (1991) (página 492) considera Pascal e binomial negativa **distintas**. A binomial negativa é definida como sendo o número de falhas antes de ocorrerem r sucessos.
  - Grinstead e Snell (1997), página 186, chamam a Pascal de binomial negativa.
  - Para Meyer(1983), página 204, a distribuição de Pascal pode ser chamada de binomial negativa.

Distribuição Poisson

▼ **Definição**

- É um modelo discreto que expressa a probabilidade de um dado número de eventos ocorrer em um intervalo contínuo (que pode ser tempo ou espaço) se esses eventos ocorrem com uma taxa média conhecida e independentemente do tempo decorrido desde o último evento.
- Seja X a v.a aleatória que conta a ocorrência de um determinado evento em um intervalo contínuo. Então X tem distribuição de Poisson com parâmetro λ e sua distribuição de probabilidade é dada por:

$$\mathbb{P}(X=x) = \frac{e^{-\lambda}\lambda^x}{x!} \quad \forall \quad x=0,1,2,\dots$$

em que, e é a base do logaritmo natural ($e=2,71828\ldots$)

- λ é um número real, igual ao número esperado de ocorrências num dado intervalo de tempo.
- **Propriedades importantes:**
  - $e^\lambda = \sum_{x=0}^{\infty}\frac{\lambda^x}{x!}$, demostramos por serie de Mac-Lourim:

$$e^{\lambda} = 1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \cdots = \sum_{x=0}^{\infty} \frac{\lambda^x}{x!}$$

- $e^{-\lambda} = 1 - \lambda + \frac{\lambda^2}{2!} - \frac{\lambda^3}{3!} + \cdots$
- $\sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda} \cdot \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{-\lambda} \cdot e^{\lambda} = 1$
- $\lim_{x\to\infty} \left(1 - \frac{\lambda}{n}\right)^{-x} = 1$

$$\lim_{x\to\infty} \left(1 - \frac{\lambda}{n}\right)^{-x} = \lim_{x\to\infty} \frac{1}{\left(1-\frac{\lambda}{n}\right)^x} = \lim_{x\to\infty} \left[\frac{1}{\left(1-\frac{\lambda}{n}\right)} \cdot \frac{1}{\left(1-\frac{\lambda}{n}\right)} \cdots \frac{1}{\left(1-\frac{\lambda}{n}\right)}\right] = \lim_{x\to\infty} \left[\frac{1}{1-0} \cdot \frac{1}{1-0} \cdots \frac{1}{1-0}\right] = 1$$

- $\lim_{x\to\infty} \left(1 - \frac{\lambda}{n}\right)^{n} = e^{-\lambda}$
- $\lim_{x\to\infty} \frac{(n)_x}{n^x} = 1$

$$\lim_{x\to\infty} \frac{(n)_x}{n^x} = \lim_{x\to\infty} \frac{n}{n} \frac{(n-1)}{n} \frac{(n-2)}{n} \cdots \frac{(n-x+1)}{n} = \lim_{x\to\infty} 1 \cdot \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{x-1}{n}\right) = (1-0)(1-0)\cdots(1-0) = 1$$

- **Demostração para $\mathbb{P}(X = x)$:**

  A variável aleatória de Poisson encontra uma tremenda faixa de aplicações em diversas áreas porque pode ser usada como uma aproximação para a variável aleatória binomial com parâmetros (n, p) no caso particular de n **grande** e p suficientemente **pequeno** para que np tenha **tamanho moderado**. Para ver isto, suponha que X seja uma variável aleatória **binomial** com parâmetros (n, p), e suponha que $\lambda = np$. Então:

$$\begin{aligned}
\mathbb{P}(X = x) &= \binom{n}{x} p^x (1-p)^{n-x} \\
&= \frac{n!}{x!(n-x)!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \qquad * p = \frac{\lambda}{n} \\
&= \frac{n(n-1)(n-2)\cdots(n-x+1)}{x!} \cdot \frac{\lambda^x}{n^x} \cdot \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x} \\
&= \frac{(n)_x}{n^x} \cdot \frac{\lambda^x}{x!} \cdot \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-x} \\
&= 1 \cdot \frac{\lambda^x}{x!} \cdot e^{-\lambda} \cdot 1 \\
&= \mathbf{\frac{e^{-\lambda}\lambda^x}{x!}}
\end{aligned}$$

- Notação: $X \sim P(\lambda)$.
  - $E(X) = \lambda$
  - $E(X^2) = \lambda(\lambda + 1)$
  - $Var(X) = \lambda$
- **Demostração da $E(X)$, $E(X^2)$, $Var(X)$**
  - ▼ $E(X) = \lambda$

$$
\begin{aligned}
\mathbb{E}(X) &= \sum_{x=0}^{\infty} x \cdot \frac{e^{-\lambda} \cdot \lambda^x}{x!} \\
&= \sum_{x=0}^{\infty} x \cdot \frac{e^{-\lambda} \cdot \lambda^{x-1} \cdot \lambda}{x(x-1)!} \\
&= \lambda e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} \\
&= \lambda e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \\
&= \lambda e^{-\lambda} \cdot e^{\lambda} \\
&= \lambda
\end{aligned}
$$

- $E(X^2) = \lambda(\lambda + 1)$

$$
\begin{aligned}
\mathbb{E}(X^2) &= \sum_{x=0}^{\infty} x^2 \cdot \frac{e^{-\lambda} \cdot \lambda^x}{x!} \\
&= \sum_{x=0}^{\infty} x^2 \cdot \frac{e^{-\lambda} \cdot \lambda^{x-1} \cdot \lambda}{x(x-1)!} \\
&= \lambda \sum_{x=0}^{\infty} \frac{x \cdot e^{-\lambda} \lambda^{x-1}}{(x-1)!} \\
&= \lambda \sum_{j=0}^{\infty} \frac{(j+1) e^{-\lambda} \lambda^j}{j!} & * j = x - 1 \\
&= \lambda \left[\sum_{j=0}^{\infty} \frac{j e^{-\lambda} \lambda^j}{j!} \right] + \left[\sum_{j=0}^{\infty} \frac{e^{-\lambda} \lambda^j}{j!} \right] \\
&= \boldsymbol{\lambda(\lambda + 1)}
\end{aligned}
$$

- $Var(X) = \lambda$

$$
Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \lambda(\lambda + 1) - \lambda^2 = \lambda^2 + \lambda - \lambda^2 = \lambda
$$

- **Exemplo:** Seja X uma v.a.d., tal que $p(x) = \frac{e^{-\lambda}\lambda^x}{x!(1-e^{-\lambda})}$ para $x = 1, 2, 3, \cdots$.

  i. Mostre que $p(x)$ é de fato uma função de probabilidade.

$$
\sum_{x=1}^{\infty} \frac{e^{-\lambda} \cdot \lambda^x}{x!(1 - e^{-\lambda})} = \frac{e^{-\lambda}}{(1 - e^{-\lambda})} \sum_{x=1}^{\infty} \frac{\lambda^x}{x!}
$$

  Note que $\sum_{x=1}^{\infty} \frac{\lambda^x}{x!}$ é igual a $\sum_{x=0}^{\infty} \frac{\lambda^x}{x!}$ menos o valor da equação quando o $x = 0$ Logo o somatório $\sum_{x=1}^{\infty} \frac{\lambda^x}{x!}$ vai ser igual a:

$$
\sum_{x=1}^{\infty} \frac{\lambda^x}{x!} = \left(\sum_{x=0}^{\infty} \frac{\lambda^x}{x!} \right) - \frac{\lambda^0}{0!} = e^{\lambda} - 1
$$

  Então,

$$
\frac{e^{-\lambda}}{(1 - e^{-\lambda})} \sum_{x=1}^{\infty} \frac{\lambda^x}{x!} = \frac{e^{-\lambda}}{(1 - e^{-\lambda})} \cdot (e^{\lambda} - 1) = \frac{e^0 - e^{-\lambda}}{(1 - e^{-\lambda})} = \frac{1 - e^{-\lambda}}{1 - e^{-\lambda}} = \mathbf{1}
$$

  ii. Determine $\mathbb{E}(X)$ e $Var(X)$.

  Para a $\mathbb{E}(X)$ temos:

$$
\begin{aligned}
\mathbb{E}(X) &= \sum_{x=1}^{\infty} x \cdot \frac{e^{-\lambda} \cdot \lambda^x}{x!(1-e^{-\lambda})} \\
&= \frac{e^{-\lambda}}{(1-e^{-\lambda})} \sum_{x=1}^{\infty} x \cdot \frac{\lambda^{(x-1)}\lambda}{x(x-1)!} \\
&= \frac{\lambda e^{-\lambda}}{(1-e^{-\lambda})} \sum_{x=1}^{\infty} \frac{\lambda^{(x-1)}}{(x-1)!} \\
&= \frac{\lambda e^{-\lambda}}{(1-e^{-\lambda})} \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} \qquad * \, y = x-1 \quad \forall y = 0,1,2,3,\cdots \\
&= \frac{\lambda e^{-\lambda}}{(1-e^{-\lambda})} \cdot e^{\lambda} \\
&= \frac{\lambda}{1-e^{-\lambda}}
\end{aligned}
$$

Para $\mathbb{E}(X^2)$ temos:

$$
\begin{aligned}
\mathbb{E}\Big[X(X-1)\Big] &= \sum_{x=1}^{\infty} x(x-1) \cdot \frac{e^{-\lambda} \cdot \lambda^x}{x!(1-e^{-\lambda})} \\
&= \frac{e^{-\lambda}}{(1-e^{-\lambda})} \sum_{x=1}^{\infty} x(x-1) \cdot \frac{\lambda^{(x-2)}\lambda^2}{x(x-1)(x-2)!} \\
&= \frac{\lambda^2 e^{-\lambda}}{(1-e^{-\lambda})} \sum_{x=1}^{\infty} \frac{\lambda^{(x-2)}}{(x-2)!} \\
&= \frac{\lambda^2 e^{-\lambda}}{(1-e^{-\lambda})} \sum_{y=0}^{\infty} \frac{\lambda^y}{y!} \qquad * \, y = x-2 \quad \forall y = 0,1,2,3,\cdots \\
&= \frac{\lambda^2 e^{-\lambda}}{(1-e^{-\lambda})} \cdot e^{\lambda} \\
&= \frac{\boldsymbol{\lambda^2}}{\mathbf{1-e^{-\lambda}}}
\end{aligned}
$$

Note que,

$$
\mathbb{E}\Big[X(X-1)\Big] = \mathbb{E}(X^2) - \mathbb{E}(X)
$$

Logo,

$$
\mathbb{E}(X^2) = \mathbb{E}\Big[X(X-1)\Big] - \mathbb{E}(X) = \frac{\boldsymbol{\lambda^2}}{\mathbf{1-e^{-\lambda}}} - \frac{\lambda}{\mathbf{1-e^{-\lambda}}} = \frac{\boldsymbol{\lambda^2+\lambda}}{\mathbf{(1-e^{-\lambda})}}
$$

Para $Var(X)$ temos:

$$
\begin{aligned}
Var(X) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \\
&= \frac{\boldsymbol{\lambda^2+\lambda}}{\mathbf{(1-e^{-\lambda})}} - \left(\frac{\lambda}{1-e^{-\lambda}}\right)^2 \\
&= \frac{\boldsymbol{\lambda^2+\lambda}}{\mathbf{(1-e^{-\lambda})}} - \frac{\lambda^2}{(1-e^{-\lambda})^2}
\end{aligned}
$$

Função Geradora de Probabilidade

▼ **Definição**

- Em teoria de probabilidade, a **Função Geradora de Probabilidade (f.g.p.)** de uma variável aleatória discreta é uma representação em **série de potências** da função de probabilidade da uma variável aleatória discreta que assume somente valores não-negativos.
- Se X for uma variável aleatória discreta assumindo valores nos inteiros não negativos $\{0, 1, \cdots\}$, então a função geradora de probabilidade de X é definida como

$$G(t) = \mathbb{E}(t^x) = \sum_x t^x \, \mathbb{P}(X = x)$$

Demostração:

$$\begin{aligned} G(t) &= P_0 t^0 + P_1 t^1 + P_2 t^2 + P_3 t^3 + \cdots + P_n t^n \\ &= P_0 + P_1 t + P_2 t^2 + P_3 t^3 + \cdots + P_n t^n \\ &= \sum_x t^x \cdot p(x) \\ &= \mathbb{E}(t^x) \end{aligned}$$

- **Ex:** Seja X uma v.a.d. com a seguinte distribuição de probabilidade.

x	0	1	2	$\cdots$	n
$p(x)$	P_0	P_1	P_2	$\cdots$	P_n

Podemos associar as probabilidades aos coeficientes de um polinômio da seguinte forma:

$$G(t) = P_0 + P_1 t + P_2 t^2 + P_3 t^3 + \cdots + P_n t^n$$

Note que para $t = 1$ temos:

$$G(1) = P_0 + P_1 + P_2 + P_3 + \cdots + P_n = 1$$

▼ **Propriedades:**

- $G(1) = P_0 + P_1 + P_2 + P_3 + \cdots + P_n = 1$
- $G(t)' = \sum_x x \; t^{x-1} p(x)$

$$G(t) = P_0 + P_1 t + P_2 t^2 + P_3 t^3 + P_4 t^4 + \cdots + P_n t^n$$

$$G(t)' = P_1 + 2P_2 t + 3P_3 t^2 + 4P_4 t^3 + \cdots + nP_n t^{n-1}$$

- $G(t)'' = \sum_x x(x-1) \; t^{x-2} p(x)$

$$G(t)'' = 2P_2 + 6P_3 t + 12P_4 t^2 + \cdots + n(n-1)P_n t^{n-2}$$

- $\mathbb{E}(X) = G(1)'$

$$\begin{gathered} G(t) = P_0 + P_1 t + P_2 t^2 + P_3 t^3 + \cdots + P_n t^n \\ G(t)' = P_1 + 2P_2 t + 3P_3 t^2 + 4P_4 t^3 \cdots + nP_n t^{n-1} \\ G(1)' = P_1 + 2P_2 + 3P_3 + 4P_4 \cdots + nP_n = \sum_{x=1}^{n} x \; p(x) = \mathbb{E}(\mathbf{X}) \end{gathered}$$

- $\mathbb{E}\Big[X(X-1)\Big] = G(1)''$

$$\begin{aligned} G(1)'' &= 2P_2 + 6P_3 \cdot 1 + 12P_4 \cdot 1^2 + \cdots + n(n-1)P_n \cdot 1^{n-2} \\ &= 2P_2 + 6P_3 + 12P_4 + \cdots + n(n-1)P_n \\ &= \sum_{x=2} x(x-1) \; \cdot \; p(x) \\ &= \mathbb{E}\Big[\mathbf{X}(\mathbf{X}-\mathbf{1})\Big] \end{aligned}$$

- $Var(X) = G(1)'' + G(1)' - \Big[G(1)'\Big]^2$

$$\begin{gathered} G(1)'' = \mathbb{E}\Big[X(X-1)\Big] = \mathbb{E}(X^2) - \mathbb{E}(X) \\ \mathbb{E}(X^2) = G(1)'' + G(1)' \; = \mathbb{E}\Big[X(X-1)\Big] + \mathbb{E}(X) = \mathbb{E}(X^2) - \mathbb{E}(X) + \mathbb{E}(X) = \mathbb{E}(X^2) \\ Var(X) = G(1)'' + G(1)' - \Big[G(1)'\Big]^2 \end{gathered}$$

- $G(0) = P_0 = p(0)$
- $G(0)' = P_1 = p(1)$
- $G(0)'' = 2P_2$
- $G(0)^{(3)} = 6P_3$
- $G(0)^{(k)} = k! P_k$
- $\frac{G(0)''}{2} = P_2 = p(2)$
- $\frac{G(0)''}{6} = P_3 = p(3)$
- $p(x) = P_x = \frac{G(0)^{(x)}}{x!}$

▼ **Binomial**

$$G(t) = \mathbb{E}(t^x) = \sum_{x=0}^{n} t^x\ p(x) = \sum_{x=0}^{n} t^x \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=0}^{n} \binom{n}{x} (tp)^x (1-p)^{n-x}$$

Note que no último somatório temos o termo geral de um binômio de newton, usando a formula $(a+b)^n = \sum_{p=0}^{n} \binom{n}{p} a^{n-p}\ b^p$, temos:

$$\sum_{x=0}^{n} \binom{n}{x} (tp)^x (1-p)^{n-x} = \Big[pt + (1-p)\Big]^n$$

Logo,

$$G(t) = \Big[pt + (1-p)\Big]^n$$

- Vamos ver agora, que as propriedades citadas anteriormente sobre função geradora, são verdadeira:
  - $G(t) = \Big[pt + (1-p)\Big]^n$
  - $G(1) = 1$

$$G(1) = \Big[p \cdot 1 + (1-p)\Big]^n = \Big[p + 1 - p\Big]^n = 1^n = 1$$

  - $G(t)' = np\Big[pt + (1-p)\Big]^{n-1}$

$$G(t)' = \left(\Big[pt + (1-p)\Big]^n\right)' = n\Big[pt + (1-p)\Big]^{n-1} \cdot (pt + 1 - p)' = n\Big[pt + (1-p)\Big]^{n-1} \cdot (p \cdot 1 + 0 - 0) = n\Big[pt + (1-p)\Big]^{n-1} \cdot p = np\Big[pt + (1-p)\Big]^{n-1}$$

  - $G(1)' = \mathbb{E}(X) = np$

$$G(1)' = np\Big[p \cdot 1 + (1-p)\Big]^{n-1} = np\Big[p + 1 - p\Big]^{n-1} = np \cdot 1^{n-1} = np$$

  - $G(t)'' = n(n-1)p^2\Big[pt + (1-p)\Big]^{n-2}$
  - $G(1)'' = n(n-1)p^2 = \mathbb{E}\big[X(X-1)\big]$

$$G(1)'' = n(n-1)p^2\Big[p \cdot 1 + (1-p)\Big]^{n-2} = n(n-1)p^2\Big[p + 1 - p\Big]^{n-2} = n(n-1)p^2(1)^{n-2} = n(n-1)p^2$$

  - $Var(X) = G(1)'' + G(1)' - \Big[G(1)'\Big]^2$

$$Var(X) = G(1)'' + G(1)' - \Big[G(1)'\Big]^2 = n(n-1)p^2 + np - (np)^2 = n^2p^2 - np^2 + np - n^2p^2 = -np^2 + np = np(1-p)$$

  - $p(0) = G(0) = (1-p)^n$
  - $p(1) = G(0)' = np(1-p)^n$
  - $p(2) = \frac{G(0)''}{2} = \frac{n(n-1)}{2} \cdot p^2(1-p)^{n-2}$
  - $p(x) = \frac{G(0)^{(x)}}{x!}$

Geométrica

$$\begin{aligned} G(t) &= \mathbb{E}(t^x) \\ &= \sum_{x=1}^{n} t^x\ p(x) \\ &= \sum_{x=1}^{n} t^x\ p(1-p)^{x-1} \\ &= p\sum_{x=1}^{n} t^x\ (1-p)^{x-1} \cdot \frac{(1-p)}{(1-p)} \\ &= \frac{p}{(1-p)} \sum_{x=1}^{n} t^x\ (1-p)^x \\ &= \frac{p}{(1-p)} \sum_{x=1}^{n} \Big[t\ (1-p)\Big]^x \\ &= \frac{p}{(1-p)} \left[t(1-p) + \Big(t(1-p)\Big)^2 + \Big(t(1-p)\Big)^3 + \cdots\right] \qquad * \text{PG } q = t(1-p) \\ &= \frac{p}{(1-p)} \cdot \frac{t(1-p)}{1 - t(1-p)} \\ &= \frac{pt}{1 - t(1-p)} \end{aligned}$$

- Note que esse resultado só é valido para $t(1-p) < 1$, logo o $t < \frac{1}{1-p}$.
- **Propriedades:**
  - $G(t) = \frac{pt}{1-t(1-p)}$
  - $G(1) = 1$

$$G(1) = \frac{p}{1-(1-p)} = \frac{p}{1-1+p} = \frac{p}{p} = 1$$

  - $G(t)' = \frac{p \cdot (1-t+pt) - pt \cdot (-1+p)}{(1-t+pt)^2} = \frac{p}{(1-t+pt)^2}$

$$G(t)' = \left(\frac{pt}{1-t(1-p)}\right)' = \left(\frac{pt}{1-t+pt}\right)' = \frac{p\cdot(1-t+pt) - pt\cdot(-1+p)}{(1-t+pt)^2} = \frac{p}{(1-t+pt)^2}$$

- $G(1)' = \mathbb{E}(X) = \frac{1}{p}$

$$G(1)' = \frac{p\cdot(1-1+p\cdot 1) - p\cdot 1\cdot(-1+p)}{(1-1+p\cdot 1)^2} = \frac{p^2+p-p^2}{p^2} = \frac{p}{p^2} = \frac{1}{p}$$

- $G(t)'' = \frac{2p-2p^2}{(1-t+pt)^3}$

$$G(t)'' = \left(\frac{p}{(1-t+pt)^2}\right)' = \frac{-p\cdot 2(1-t+pt)(-1+p)}{(1-t+pt)^4} = -\frac{2p(-1+p)}{(1-t+pt)^3} = -\frac{-2p+2p^2}{(1-t+pt)^3} = \frac{2p-2p^2}{(1-t+pt)^3}$$

- $G(1)'' = \mathbb{E}\Big[X(X-1)\Big] = \frac{2-2p}{p^2}$

$$G(1)'' = \frac{2p-2p^2}{(1-1+p\cdot 1)^3} = \frac{2p-2p^2}{p^3} = \frac{p(2-2p)}{p^3} = \frac{2-2p}{p^2}$$

- $Var(X) = G(1)'' + G(1)' - \Big[G(1)'\Big]^2 = \frac{1-p}{p^2}$

$$Var(X) = G(1)'' + G(1)' - \Big[G(1)'\Big]^2 = \frac{2-2p}{p^2} + \frac{1}{p} - \left[\frac{1}{p}\right]^2 = \frac{2-2p}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$$

▼ **Poisson**

$$\begin{aligned} G(t) &= \mathbb{E}(t^x) \\ &= \sum_{x=1}^{n} t^x\ p(x) \\ &= \sum_{x=1}^{n} t^x\ \frac{e^{-\lambda}\cdot\lambda^x}{x!} \\ &= e^{-\lambda}\sum_{x=1}^{n}\frac{(t\cdot\lambda)^x}{x!} \\ &= e^{-\lambda}\cdot e^{t\lambda} \\ &= e^{-\lambda+t\lambda} \\ &= e^{\lambda(t-1)} \end{aligned}$$

▼ **Propriedades:**

- $G(t) = e^{\lambda(t-1)}$
- $G(1) = 1$

$$G(1) = e^{\lambda(1-1)} = e^0 = 1$$

- $G(t)' = \lambda e^{\lambda t-\lambda}$

$$G(t)' = \left(e^{\lambda(t-1)}\right)' = \left(e^{\lambda t-\lambda}\right)' = e^{\lambda t-\lambda}\cdot(\lambda t-\lambda)' = e^{\lambda t-\lambda}\cdot(\lambda-0) = \lambda e^{\lambda t-\lambda}$$

- $G(1)' = \mathbb{E}(X) = \lambda$

$$G(1)' = \lambda e^{\lambda-\lambda} = \lambda e^0 = \lambda$$

- $G(t)'' = \lambda^2 e^{\lambda t-\lambda}$

$$G(t)' = \left(\lambda e^{\lambda t-\lambda}\right)' = \lambda e^{\lambda t-\lambda}\cdot(\lambda t-\lambda)' = \lambda e^{\lambda t-\lambda}\cdot\lambda = \lambda^2 e^{\lambda t-\lambda}$$

- $G(1)'' = \mathbb{E}\Big[X(X-1)\Big] = \lambda^2$

$$G(1)'' = \lambda^2 e^{\lambda-\lambda} = \lambda^2 e^0 = \lambda^2$$

- $Var(X) = G(1)'' + G(1)' - \Big[G(1)'\Big]^2 = \lambda$

$$Var(X) = G(1)'' + G(1)' - \Big[G(1)'\Big]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

▼ **Exemplos**

▼ **Ex1:** Considere X uma v.a.d assumindo valores não-negativos com a seguinte f.g.p:

$$G_X(t) = A\left(\frac{10+8t^2}{2-t}\right) \qquad \forall t \in (-2, 2)$$

a. Determine o valor de A.

Sabemos que $G_X(1) = 1$, logo

$$G_X(1) = A\left(\frac{10+8\cdot 1^2}{2-1}\right) = A(10+8) = 1$$

$$A(10+8) = 1 \iff A = \frac{1}{18}$$

b. Determine o valor esperado de X.

Sabemos que o $\mathbb{E}(X) = G(1)'$, logo

$$G_X(t)' = \left[\frac{1}{18}\cdot\left(\frac{10+8t^2}{2-t}\right)\right]' = \frac{1}{18}\cdot\left(\frac{10+8t^2}{2-t}\right)' = \frac{1}{18}\cdot\left(\frac{16t\cdot(2-t)-(10+8t^2)\cdot(-1)}{(2-t)^2}\right) = \frac{1}{18}\cdot\left(\frac{32t-8t^2+10}{(2-t)^2}\right)$$

Para $G_X(1)'$ temos:

$$G_X(1)' = \frac{1}{18}\cdot\left(\frac{32-8+10}{(2-1)^2}\right) = \frac{1}{18}\cdot\frac{34}{1} = \frac{34}{18} = \frac{17}{9}$$
$$\mathbb{E}(X) = \frac{17}{9}$$

▼ **Ex2:** Se se $X \sim Poisson(\lambda)$ prove com f.g.p que:

$$\mathbb{E}(X_{(r)}) = \mathbb{E}\Big[X(X-1)(X-2)\cdots(X-r+1)\Big] = \lambda^r$$

Vamos começar provando que $\mathbb{E}(X_{(r)}) = \mathbb{E}\Big[X(X-1)(X-2)\cdots(X-r+1)\Big]$

$$\begin{aligned}
G'(t) &= \sum_x xt^{x-1}p(x)\\
G'(1) &= \sum_x xp(x) = \mathbb{E}(X)\\
G''(t) &= \sum_x x(x-1)t^{x-2}p(x)\\
G''(1) &= \sum_x x(x-1)p(x) = \mathbb{E}\Big[X(X-1)\Big] = \mathbb{E}(X_{(2)})\\
G'''(t) &= \sum_x x(x-1)(x-2)t^{x-3}p(x)\\
G'''(1) &= \sum_x x(x-1)(x-2)p(x) = \mathbb{E}\Big[X(X-1)(X-2)\Big] = \mathbb{E}(X_{(3)})\\
G^{(r)}(t) &= \sum_x x(x-1)(x-2)\cdots(x-r+1)t^{x-r}p(x)\\
G^{(r)}(1) &= \sum_x x(x-1)(x-2)\cdots(x-r+1)p(x) = \mathbb{E}\Big[X(X-1)(X-2)\cdots(x-r+1)\Big] = \mathbb{E}(X_{(r)})
\end{aligned}$$

Agora vamos prova que $\mathbb{E}(X_{(r)}) = \lambda^r$, note que:

$$G(t) = \mathbb{E}(t^X) = \sum_{x=0}^{\infty} t^x\frac{e^{-\lambda}\cdot\lambda^x}{x!} = e^{-\lambda}\cdot\sum_{x=0}^{\infty}\frac{(t\lambda)^x}{x!} = e^{-\lambda}\cdot e^{t\lambda} = e^{\lambda(t-1)}$$
$$\begin{aligned}
G'(t) &= \lambda e^{\lambda(t-1)} \Rightarrow G'(1) = \lambda\\
G''(t) &= \lambda^2 e^{\lambda(t-1)} \Rightarrow G''(1) = \lambda^2\\
G'''(t) &= \lambda^3 e^{\lambda(t-1)} \Rightarrow G'''(1) = \lambda^3\\
&\vdots\\
G^{(r)}(t) &= \lambda^r e^{\lambda(t-1)} \Rightarrow G^{(r)}(1) = \lambda^r
\end{aligned}$$

Note que $\mathbb{E}(X_{(r)}) = G^{(r)}(1) = \lambda^r$.

Função Geradora de Momento

▼ **Definição**

- Em teoria de probabilidade a função geradora de momento (**f.g.m**) de uma v.a. discreta ou continua fornece uma especificação alternativa para a distribuição de probabilidade.
- **Definição:** Seja X uma v.a.d ou v.a.c. A função geradora de momentos de X é definida por:

$$M_X(t) = \mathbb{E}(e^{tX})$$

deste que a esperança seja finita para $t \in \mathbb{R}$ em algum inteiro $-t_0 < t < t_0$ com $t_0 > 0$.

Obs: Se essa condição não for satisfeita, dizemos que a f.g.m não existe.

Obs: No caso de X ser uma v.a.d, então:

$$M_X(t) = \mathbb{E}(e^{tX}) = \sum_x e^{tx}p(x)$$

Note que:

$$\begin{aligned}
e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots\\
e^{tx} &= 1 + tx + \frac{(tx)^2}{2!} + \frac{(tx)^3}{3!} + \cdots
\end{aligned}$$

Logo,

$$M_x(t) = \mathbb{E}(e^{tX}) = 1 + t\mathbb{E}(X) + \frac{t^2\mathbb{E}(X^2)}{2!} + \frac{t^3\mathbb{E}(X^3)}{3!} + \cdots$$

- Chamamos $M(t)$ de função geradora de momentos porque todos os momentos de X podem ser obtidos com o cálculo sucessivo da derivada de $M(t)$ e então com sua avaliação em $t = O$.
- A função geradora de momentos determina de forma unívoca a distribuição da variável aleatória, ou seja, existe uma única distribuição com função geradora $m(t)$. Por outro lado, se duas variáveis aleatórias possuem uma mesma função geradora então possuem a mesma distribuição.

▼ **Propriedades:**

- $M_X(0) = \mathbb{E}(e^{0X}) = \mathbb{E}(1) = 1$
- ▼ $M'(t) = \mathbb{E}(X) + t\mathbb{E}(X^2) + \frac{t^2\mathbb{E}(X^3)}{2} + \frac{t^3\mathbb{E}(X^4)}{6} + \cdots$

$$M'(t) = \mathbb{E}(e^{tX})' = \left(1 + t\mathbb{E}(x) + \frac{t^2\mathbb{E}(X^2)}{2!} + \frac{t^3\mathbb{E}(X^3)}{3!} + \frac{t^4\mathbb{E}(X^4)}{4!} + \cdots\right)'$$
$$= 0 + \mathbb{E}(x) + \frac{2t\mathbb{E}(X^2)}{2!} + \frac{3t^2\mathbb{E}(X^3)}{3!} + \frac{4t^3\mathbb{E}(X^4)}{4!} + \cdots$$
$$= \mathbb{E}(x) + t\mathbb{E}(X^2) + \frac{t^2\mathbb{E}(X^3)}{2} + \frac{t^3\mathbb{E}(X^4)}{6} + \cdots$$

- $M'(0) = \mathbb{E}(X)$
- ▼ $M''(t) = \mathbb{E}(X^2) + t\mathbb{E}(X^3) + \frac{t^2\mathbb{E}(X^4)}{2} + \frac{t^3\mathbb{E}(X^5)}{6} + \cdots$

$$M''(t) = \mathbb{E}(e^{tX})'' = \left(\mathbb{E}(x) + t\mathbb{E}(X^2) + \frac{t^2\mathbb{E}(X^3)}{2} + \frac{t^3\mathbb{E}(X^4)}{6} + \frac{t^4\mathbb{E}(X^5)}{24} + \cdots\right)'$$
$$= 0 + \mathbb{E}(X^2) + \frac{2t\mathbb{E}(X^3)}{2} + \frac{3t^2\mathbb{E}(X^4)}{6} + \frac{4t^3\mathbb{E}(X^5)}{24} + \cdots$$
$$= \mathbb{E}(X^2) + t\mathbb{E}(X^3) + \frac{t^2\mathbb{E}(X^4)}{2} + \frac{t^3\mathbb{E}(X^5)}{6} + \cdots$$

- $M''(0) = \mathbb{E}(X^2)$
- $M^{(r)}(t) = \mathbb{E}(X^r) + t\mathbb{E}(X^{r+1}) + \sum_{n=2}^{\infty} \frac{t^n \cdot \mathbb{E}(X^{r+n})}{n!}$
- $M^{(r)}(0) = \mathbb{E}(X^r)$
- $M''(0) - [M'(0)]^2 = Var(X)$

▼ **Função Característica**

- Existe uma outra função, estreitamente relacionada com a f.g.m, a qual é frequentemente empregada em seu lugar.
- Ela é denominada função característica e é denota por:

$$\phi_X(t) = \mathbb{E}(e^{itX}) = \mathbb{E}[\cos(tX)] + i\ \mathbb{E}[\sin(tX)]$$

para $t \in \mathbb{R}$ e $i = \sqrt{-1}$ (unidade imaginária).

- As funções característica tem a vantagem de sempre existir. Entretanto, teremos o inconveniente de trabalhar com uma função de valores complexos.

▼ **Binomial**

$$M_X(t) = \mathbb{E}(e^{tX}) = \sum_{x=0}^{n} e^{tx}\binom{n}{x}p^x(1-p)^{n-x}$$
$$= \sum_{x=0}^{n}\binom{n}{x}(e^t p)^x(1-p)^{n-x}$$
$$= (pe^t + 1 - p)^n$$

- A última igualdade decorre de uma aplicação direta do teorema binomial.

▼ **Propriedades:**

- $M'(t) = n(pe^t + 1 - p)^{n-1}pe^t$
- $M'(0) = \mathbb{E}(X) = np$

$$\mathbb{E}(X) = M'(0) = n(pe^0 + 1 - p)^{n-1}pe^0 = n(p + 1 - p)^{n-1}p = np(1)^{n-1} = np$$

- $M''(t) = npe^t \cdot (pe^t + 1 - p)^{n-1} + (n^2p^2e^{2t} - np^2e^{2t}) \cdot (pe^t + 1 - p)^{n-2}$

$$M''(t) = \left(n(pe^t + 1 - p)^{n-1}pe^t\right)'$$
$$= \left(npe^t(pe^t + 1 - p)^{n-1}\right)'$$
$$= \left((npe^t)' \cdot (pe^t + 1 - p)^{n-1}\right) + \left(npe^t \cdot \left((pe^t + 1 - p)^{n-1}\right)'\right)$$
$$= \left(npe^t \cdot (pe^t + 1 - p)^{n-1}\right) + \left(npe^t \cdot (n-1)(pe^t + 1 - p)^{n-2} \cdot pe^t\right)$$
$$= \left(npe^t \cdot (pe^t + 1 - p)^{n-1}\right) + \left((n^2p^2e^{2t} - np^2e^{2t}) \cdot (pe^t + 1 - p)^{n-2}\right)$$
$$= npe^t \cdot (pe^t + 1 - p)^{n-1} + (n^2p^2e^{2t} - np^2e^{2t}) \cdot (pe^t + 1 - p)^{n-2}$$

- $M''(0) = \mathbb{E}(X^2) = n^2p^2 + np(1-p)$

$$M''(0) = npe^0 \cdot (pe^0 + 1 - p)^{n-1} + (n^2p^2e^{2\cdot 0} - np^2e^{2\cdot 0}) \cdot (pe^0 + 1 - p)^{n-2}$$
$$= np \cdot (p + 1 - p)^{n-1} + (n^2p^2 - np^2) \cdot (p + 1 - p)^{n-2}$$
$$= np \cdot 1^{n-1} + (n^2p^2 - np^2) \cdot 1^{n-2}$$
$$= np + n^2p^2 - np^2$$
$$= n^2p^2 + np(1-p)$$

- $Var(X) = M(0)'' - \left[M(0)'\right]^2$

$$Var(X) = M(0)'' - \left[M(0)'\right]^2 = n^2p^2 + np(1-p) - (np)^2 = np(1-p)$$

▼ **Geométrica**

$$
\begin{aligned}
M_X(t) = \mathbb{E}(e^{tX}) &= \sum_{x=1}^{\infty} e^{tx} p(1-p)^{x-1} \\
&= \sum_{x=1}^{\infty} e^{tx} p(1-p)^{x-1} \cdot \frac{(1-p)}{1-p} \\
&= \frac{p}{1-p} \sum_{x=1}^{\infty} \left[e^t(1-p)\right]^x \\
&= \frac{p}{q} \sum_{x=1}^{\infty} (e^t q)^x \\
&= \frac{p}{q} \left[e^t q + (e^t q)^2 + (e^t q)^3 + \cdots\right] \\
&= \frac{p}{q} \cdot \frac{e^t q}{1 - e^t q} \\
&= \frac{pe^t}{1 - e^t q}
\end{aligned}
$$

Note que essa resultado só é válido para $\forall \;\; e^t q < 1$, logo $e^t < \frac{1}{q} \implies t < \ln(\frac{1}{q})$.

▼ **Propriedades:**

- $M(0) = 1$

$$
M(0) = \frac{p \cdot e^0}{1 - e^0 q} = \frac{p}{1-q} = \frac{p}{1-(1-p)} = \frac{p}{p} = 1
$$

- $M'(t) = \frac{pe^t}{(1-e^t q)^2}$

$$
M(t)' = \frac{pe^t \cdot (1 - e^t q) - pe^t \cdot e^t q}{(1 - e^t q)^2} = \frac{pe^t - pqe^{2t} - pqe^{2t}}{(1 - e^t q)^2} = \frac{pe^t}{(1 - e^t q)^2}
$$

- $M'(0) = \mathbb{E}(X) = \frac{1}{p}$

$$
\mathbb{E}(X) = M'(0) = \frac{pe^0}{(1 - e^0 q)^2} = \frac{p}{(1-q)^2} = \frac{p}{\left(1 - (1-p)\right)^2} = \frac{p}{p^2} = \frac{1}{p}
$$

▼ $M''(t) = \frac{pe^t(1+qe^t)}{(1-e^t q)^3}$

$$
\begin{aligned}
M(t)'' &= \left(\frac{pe^t}{(1 - e^t q)^2}\right)' \\
&= \frac{pe^t(1 - e^t q)^2 - pe^t 2(1 - e^t q)(-e^t q)}{(1 - e^t q)^4} \\
&= \frac{pe^t(1 - e^t q)^2 + 2pqe^{2t}(1 - e^t q)}{(1 - e^t q)^4} \\
&= \frac{(1 - e^t q)\left[pe^t(1 - e^t q) + 2pqe^{2t}\right]}{(1 - e^t q)^4} \\
&= \frac{pe^t(1 - e^t q) + 2pqe^{2t}}{(1 - e^t q)^3} \\
&= \frac{pe^t - pqe^{2t} + 2pqe^{2t}}{(1 - e^t q)^3} \\
&= \frac{pe^t + pqe^{2t}}{(1 - e^t q)^3} \\
&= \frac{pe^t(1 + qe^t)}{(1 - e^t q)^3}
\end{aligned}
$$

▼ $M''(0) = \mathbb{E}(X^2) = \frac{2-p}{p^2}$

$$
\begin{aligned}
M''(0) &= \frac{pe^0(1 + qe^0)}{(1 - e^0 q)^3} \\
&= \frac{p(1+q)}{(1-q)^3} \\
&= \frac{p(1 + (1-p))}{(1 - (1-p))^3} \\
&= \frac{p(2-p)}{p^3} \\
&= \frac{2-p}{p^2}
\end{aligned}
$$

- $Var(X) = M(0)'' - \left[M(0)'\right]^2$

$$
Var(X) = M(0)'' - \left[M(0)'\right]^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}
$$

▼ **Poisson**

$$
\begin{aligned}
M_X(t) = \mathbb{E}(e^{tX}) &= \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda} \cdot \lambda^x}{x!} \\
&= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} \\
&= e^{-\lambda} \cdot e^{\lambda e^t} \\
&= e^{-\lambda + \lambda e^t} \\
&= e^{\lambda(e^t - 1)}
\end{aligned}
$$

- **Exemplos**
  - **Ex1:** Se $X \sim B(n, p)$ determine a distribuição de $Y = n - X$ utilizando a f.g.m de X.

$$
M_Y(t) = \mathbb{E}(e^{tY}) = \mathbb{E}(e^{t(n-X)}) = \mathbb{E}(e^{tn} \cdot e^{-tX}) = e^{tn} \cdot \mathbb{E}(e^{-tX}) = e^{tn} \cdot M_X(-t)
$$

Como v.a.d X tem distribuição binomial com parâmetros n e p, e sabemos que a f.g.m de uma distribuição binomial é $(pe^t + 1 - p)^n$, logo temos:

$$
\begin{aligned}
e^{tn} \cdot M_X(-t) = e^{tn} \cdot (pe^{-t} + 1 - p)^n = (e^t)^n \cdot (pe^{-t} + 1 - p)^n = \left[(e^t) \cdot (pe^{-t} + 1 - p)\right]^n = \left[(e^t) \cdot (pe^{-t} + 1 - p)\right]^n \\
= \left[p + e^t(1 - p))\right]^n
\end{aligned}
$$

Então Y tem distribuição binomial com parâmetros n e $(1 - p)$.

$Y \sim B(n; 1 - p)$

  - **Ex2:** Se X. tem f.g.m $M_X(t)$, determine a f.g.m de $Y = aX + b$.

$$
M_Y(t) = \mathbb{E}(e^{tY}) = \mathbb{E}(e^{t(aX+b)}) = \mathbb{E}(e^{taX} \cdot e^{tb}) = e^{bt} \cdot \mathbb{E}(e^{atX}) = e^{bt} \cdot M_X(at)
$$

  - **Ex3:** Seja X uma v.a. tal que $p(x) = \frac{e^{-\lambda}\lambda^x}{x!(1-e^{-\lambda})} \quad \forall x = 1, 2, 3 \cdots$

a. Determine o f.g.m de X.

$$
\begin{aligned}
M_X(t) = \mathbb{E}(e^{tX}) &= \sum_{x=1}^{\infty} e^{tx} \frac{e^{-\lambda} \cdot \lambda^x}{x!(1 - e^{-\lambda})} \\
&= \frac{e^{-\lambda}}{(1 - e^{-\lambda})} \sum_{x=1}^{\infty} \frac{(\lambda e^t)^x}{x!} + 1 - 1 \\
&= \frac{e^{-\lambda}}{(1 - e^{-\lambda})} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} - 1 \\
&= \frac{e^{-\lambda}}{(1 - e^{-\lambda})} \cdot (e^{\lambda e^t} - 1) \\
&= \frac{e^{\lambda e^t - \lambda} - e^{-\lambda}}{(1 - e^{-\lambda})} \\
&= \frac{e^{\lambda e^t - \lambda} - \frac{1}{e^\lambda}}{1 - \frac{1}{e^\lambda}} \\
&= \frac{\frac{e^{\lambda e^t} - 1}{e^\lambda}}{\frac{e^\lambda - 1}{e^\lambda}} \\
&= \frac{e^{\lambda e^t} - 1}{e^\lambda - 1}
\end{aligned}
$$

b. Determine o valor esperado de X usando a f.g.m.

$$
\begin{aligned}
M_X(t)' &= \left(\frac{e^{\lambda e^t} - 1}{e^\lambda - 1}\right)' \\
&= \frac{1}{e^\lambda - 1}\left(e^{\lambda e^t} - 1\right)' \\
&= \frac{1}{e^\lambda - 1}\left(e^{\lambda e^t} \cdot \lambda e^t\right) \\
&= \frac{e^{\lambda e^t} \cdot \lambda e^t}{e^\lambda - 1}
\end{aligned}
$$

Logo,

$$
\mathbb{E}(X) = M_X(0)' = \frac{e^{\lambda e^0} \cdot \lambda e^0}{e^\lambda - 1} = \frac{e^\lambda \cdot \lambda}{e^\lambda - 1}
$$