Introdução aos Processos Estocásticos - Variáveis Aleatórias

Eduardo M. A. M. Mendes

DELT - UFMG
Programa de Pós-Graduação em Engenharia Elétrica
Universidade Federal de Minas Gerais
emmendes@cpdee.ufmg.br

Idéia: Uma variável aleatória é uma função definida no espaço amostral.

- Fenômeno aleatório numericamente válido. Exemplo: Dados

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

- Normalmente é necessário algum tipo de mapa.

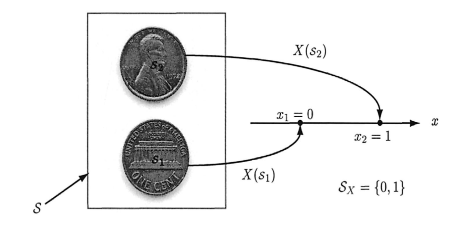


Figura 1: Exemplo de Mapa - Variável Aleatória

Variável Aleatória

Na sucessão de M lançamentos da moeda, o interesse é saber, por exemplo, o número total de caras observadas:

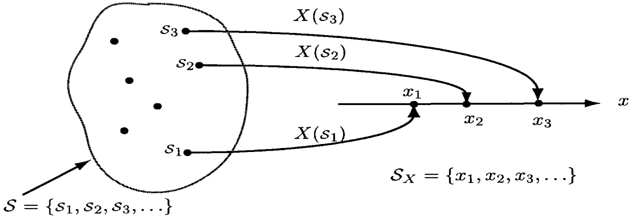
$$X(s_i) = \begin{cases} 0 & s_i = \text{cara} \\ 1 & s_i = \text{coroa} \end{cases}$$

e o número total de caras

$$\sum_{i=1}^{M} X(s_i)$$

A função que mapeia Ω em Ω_x é chamada variável aleatória.

Importante: O mapeamento $X(\bullet)$ não é aleatório, pelo contrário é normalmente conhecido e escolhido. O que é aleatório é s_i.



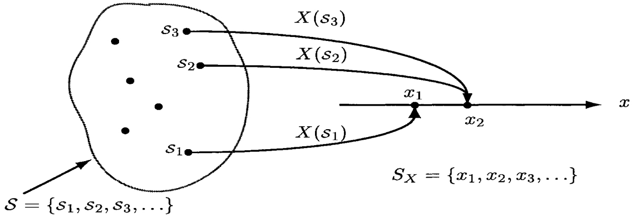


Figura 2: Tipos de Mapas

Variável Aleatória

Em geral uma variável aleatória é uma função que mapeia o espaço amostral Ω em um subconjunto em $\mathbb{R}$.

Para uma variável discreta, esse subconjunto é um conjunto finito ou enumerável de pontos.

O subconjunto forma um novo espaço amostral Ω_x

Exemplo 6.1 - Morettin, página 129

Empresário - firma - produto

$$\text{produto} \begin{cases} \text{esfera} \\ \text{cilindro} \end{cases}$$

Variável Aleatória (cont.)

- Esfera - adquirida na fábrica A
- Cilindro - adquirido na fábrica B
- Produto - comprimento definido pelo cilindro e espessura pela esfera - Verificado após montagem.

Objetivo: Viabilidade - distribuição de lucro por peça montada

$$\text{Cada componente (R\$5,00)} \begin{cases} B & - & \text{Bom} \\ L & - & \text{Longo} \\ C & - & \text{Curto} \end{cases}$$

- Se o produto final tem um C - sucata por R\$5,00
- Cada componente longo pode ser recuperado por R\$5,00

Produto	Fáb. A (Cilindro)	Fáb. B (Esfera)
Dentro das especificações - B	$0,80$	$0,70$
Maior - L	$0,10$	$0,20$
Menor - C	$0,10$	$0,10$

Preço de venda R$$25,00$ - distribuição de frequências da variável X.

Lucro por conjunto montado?

Solução

- Espaço amostral - considerando que a classificação dos cilindros e da esfera são eventos independentes e usando o princípio multiplicativo, temos o seguinte diagrama representativo das probabilidades envolvidas

Variável Aleatória (cont.)

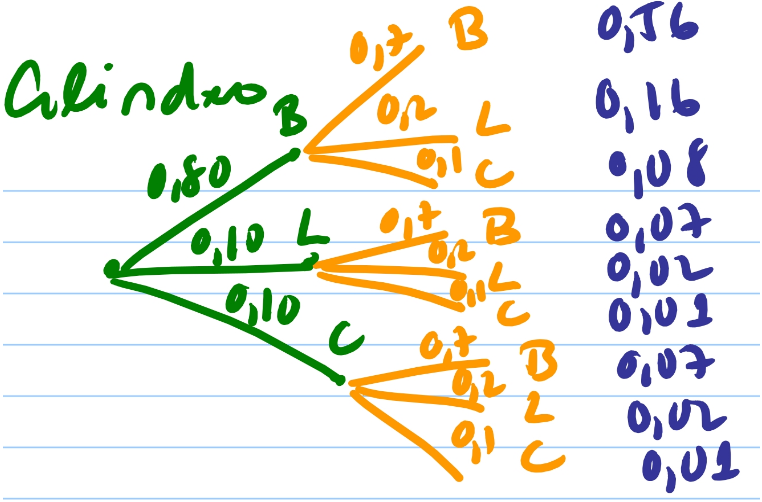


Figura 3: Eventos e Probabilidades

Variáveis Aleatórias

Produto	Probabilidade	Lucro por Montagem (X)
BB	$0,56$	$15\ (25-5-5)$
BL	$0,16$	$10\ (25-5-5-5)$
BC	$0,08$	-5
LB	$0,07$	$10\ (25-5-5-5)$
LL	$0,02$	$5\ (25-5-5-5-5)$
LC	$0,01$	-5
CB	$0,07$	-5
CL	$0,02$	-5
CC	$0,01$	-5

Vamos associar os valores:

$$\begin{array}{rclcrcl} 15 & \rightarrow & A_1=\{BB\} & \rightarrow & P(A_1) & = & 0,56 \\ 10 & \rightarrow & A_2=\{BL,LB\} & \rightarrow & P(A_2) & = & 0,16+0,07=0,23 \\ 5 & \rightarrow & A_3=\{LL\} & \rightarrow & P(A_3) & = & 0,02 \\ -5 & \rightarrow & A_4=\{BC,LC,CB,CL,CC\} & \rightarrow & P(A_4) & = & 0,08+0,01+0,07+ \\ & & & & & & 0,02+0,01=0,19 \end{array}$$

Podemos, então, escrever um modelo teórico para distribuição de X

x	$p(x)$
15	0,56
10	0,23
5	0,02
-5	0,19
	1,00

A função $(x, p(x))$ é chamada função de probabilidade da variável X.

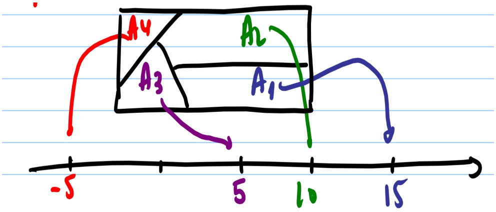


Figura 4: Mapa - Variável Aleatória X - Cilindro e Esfera

Se considerarmos Y como sendo a variável "custo de recuperação de cada produto produzido", teremos

$$\begin{array}{lcl} 0 & \rightarrow & B_1 = \{BB, BC, LC, CB, CL, CL, CC\} \\ 5 & \rightarrow & B_2 = \{BL, LB\} \\ 10 & \rightarrow & B_3 = \{LL\} \end{array}$$

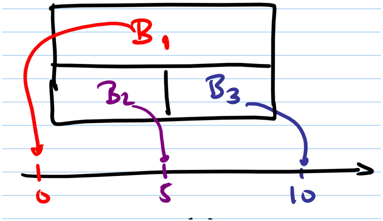


Figura 5: Mapa - Variável Aleatória Y - Cilindro e Esfera

x	$p(x)$
0	0,75
5	0,23
10	0,12
	1,00

Definição

Uma função X definida no espaço amostral Ω e com valores em um conjunto enumerável de pontos da reta é dita variável aleatória discreta.

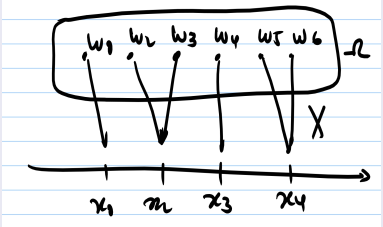


Figura 6: Ilustração da definição de variável aleatória

- Probabilidade - $P[X(S_i) = x_i] = P[\{s(i)\}]$
- Função Massa de Probabilidade (PMF) - $p_X[x_i] = P[X(s_i) = x_i]$

Propriedades

Propriedade 1

$$0 \leq p_X[x_i] \leq 1$$

Propriedade 2

$$\begin{aligned} \sum_{i=1}^{M} p_X[x_i] &= 1 \text{ se } \Omega_X \text{ consiste de M resultados} \\ \sum_{i=1}^{\infty} p_X[x_i] &= 1 \text{ se } \Omega_X \text{ enumerável} \end{aligned}$$

Para um evento A definido em Ω_X, a probabilidade é dada por

$$P[X \in A] = \sum_{\{i = x_i \in A\}} p_X[x_i]$$

Função Massa de Probabilidade

Bernoulli

Experimentos cujos resultados apresentam ou não uma característica

$$p_X[k] = \begin{cases} 1-p, & k=0 \\ p, & k=1 \end{cases}$$

Notação: $Ber(p)$

$$X \quad \sim \quad Ber(p)$$

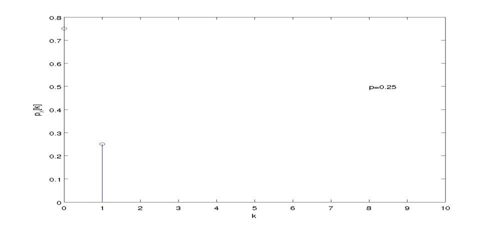


Figura 7: PMF Bernoulli.

Função Massa de Probabilidade

Binomial

- n ensaios de Bernoulli
- ensaios independentes
- probabilidade de sucesso em cada ensaio é p.

$$p_X[k] = \binom{M}{k} p^k (1-p)^{M-k}, \; k = 0, 1, \ldots, M$$

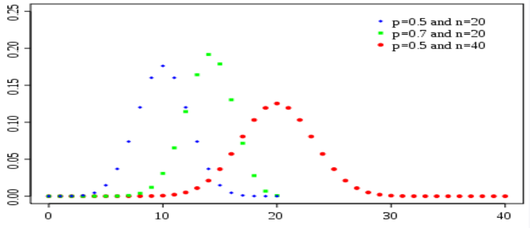


Figura 8: PMF Binomial. O máximo é localizado em $[(M+1)p]$

Função Massa de Probabilidade

Geométrica

$$p_X[k] = (1-p)^{k-1}p, \;\; k = 1, 2, \ldots$$

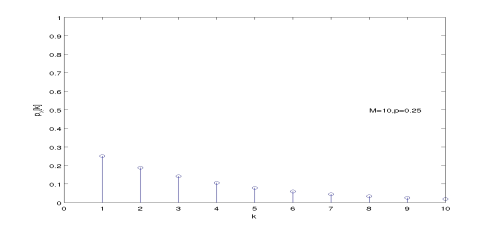


Figura 9: PMF Geométrica

Função Massa de Probabilidade

Poisson

$$p_X[k] = e^{-k}\frac{\lambda^k}{k!},\ \ k = 0, 1, 2, 3 \ldots$$

$Pois(\lambda)$ $\begin{cases} \text{Quando se deseja contar o número de eventos de um certo} \\ \text{tipo em um intervalo de tempo, superfície ou volume} \end{cases}$

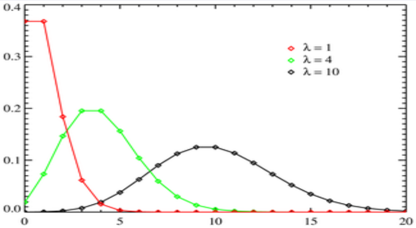


Figura 10: PMF Poisson

Função Massa de Probabilidade

Uniforme Discreta

$$p_X[k] = p, \;\; k = 1, \ldots, N$$

onde $p = \frac{1}{N}$.

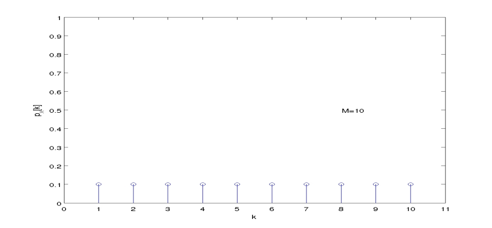


Figura 11: PMF Uniforme Discreta

Função Massa de Probabilidade

Hipergeométrica

$$p_X[k] = \frac{\binom{r}{k}\binom{N-r}{n-k}}{\binom{N}{n}}$$

com $0 \leq k \leq min(r, n)$

$$\text{Sem reposição dividida em dois atributos} \begin{cases} N \text{ objetivos} \\ r \text{ tem atributo } A \\ N - r \text{ tem atributo } B \\ n \text{ retirada} \end{cases}$$

Objetivo: Determinar a PMF de $Y = g(X)$, onde X é uma variável aleatória discreta.

Exemplo kay página 115

Dados com faces

$$0, 0, 1, 1, 2, 2$$

Achar a PMF do número observado quando o dado é lançado, assumindo que todos os lados têm probabilidade igual de ocorrer.

Transformação de Variáveis Aleatórias Discretos (cont.)

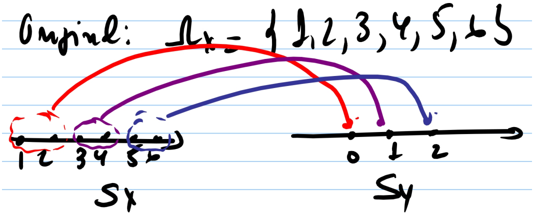


Figura 12: Exemplo Kay

$$
y = \begin{cases} y_1 = 0, & \text{se } x = x_1 = 1 \text{ ou } x = x_2 = 2 \\ y_2 = 1, & \text{se } x = x_3 = 3 \text{ ou } x = x_4 = 4 \\ y_3 = 2, & \text{se } x = x_5 = 5 \text{ ou } x = x_6 = 6 \end{cases}
$$

$$p_Y[y_i] = \begin{cases} p_X[1] + p_X[2] & = & \frac{1}{3}, \; i = 1 \\ p_X[3] + p_X[4] & = & \frac{1}{3}, \; i = 2 \\ p_X[5] + p_X[6] & = & \frac{1}{3}, \; i = 3 \end{cases}$$

Em geral, temos

$$p_Y[y_i] = \sum_{\{j = g(x_j) = y_i\}} p_X[x_j]$$

Transformação VAs - Exemplo 1

$$\begin{aligned} X &\sim Ber(p) \\ Y &= 2X - 1? \end{aligned}$$

Solução:

$$\begin{aligned} \Omega_x &= \{0, 1\} \\ &\downarrow \\ \Omega_y &= \{-1, 1\} \end{aligned}$$

$$\begin{aligned} x_1 = 0 &\rightarrow y_1 = -1 \\ x_2 = 1 &\rightarrow y_2 = 1 \end{aligned}$$

Transformação VAs - Exemplo 1 (cont.)

$$
\begin{aligned}
p_Y[-1] &= p_X[0] = 1 - p \\
p_Y[1] &= p_X[1] = p
\end{aligned}
$$

$$\begin{aligned} Y &= g(X) = X^2 \\ \Omega &= \{-1, 0, 1\} \end{aligned}$$

Solução:

$$\begin{cases} x_1 = -1 \rightarrow y_1 = 1 \\ x_2 = 0 \rightarrow y_2 = 0 \\ x_3 = 1 \rightarrow y_3 = 1 \end{cases}$$

As probabilidades são:

$$\begin{aligned} p_Y[1] &= p_X[-1] + p_X[1] \\ p_Y[0] &= p_X[0] \end{aligned}$$

onde $\Omega_X \rightarrow \Omega_Y = \{0, 1\}$

Função de Distribuição Acumulada (CDF)

Definição

Dada uma variável aleatória X, a função de distribuição de acumulada (ou simplesmente função de distribuição) é a função

$$F_X(x) = p[X \leq x], \qquad -\infty < x < +\infty$$

Função de Distribuição Acumulada (CDF)

Exemplo Binomial

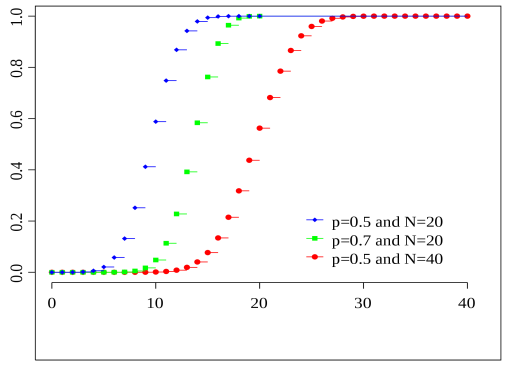


Figura 13: Exemplo CDF Binomial

Função de Distribuição Acumulada (CDF)

Exemplo Geométrica

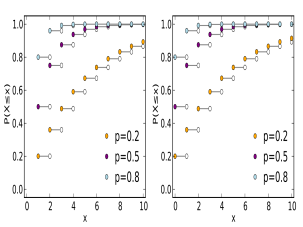


Figura 14: Exemplo CDF Geométrica

- É possível recuperar a $p_X[x]$ da $F_X[x]$

$$p_X[x] = F_X[x^+] - F_X[x^-]$$

onde x^+ é um ponto um pouco acima de x e x^- é um ponto um pouco abaixo de x.

CDF

Propriedade 1

A CDF é um número entre 0 e 1.

$$0 \leq F_X[x] \leq 1$$

Prova: Usando a definição

$$F_X[x] = P[X \leq x]$$

O segundo lado da igualdade é uma probabilidade e, portanto, entre 0 e 1. □

CDF

Propriedade 2

Os Limites da CDF quando $x \to -\infty$ e $x \to +\infty$ são:

$$\begin{aligned}
\lim_{x \to -\infty} F_X[x] &= 0 \\
\lim_{x \to +\infty} F_X[x] &= 1
\end{aligned}$$

Prova:

$$\begin{aligned}
\lim_{x \to -\infty} F_X[x] &= P[\{s : X(s) < -\infty\}] &= P[\emptyset] &= 0 \\
\lim_{x \to -\infty} F_X[x] &= P[\{s : X(s) < +\infty\}] &= P[\Omega] &= 1
\end{aligned}$$

CDF

Propriedade 3

A CDF é uma função monotonicamente crescente.

Prova:

$$
\begin{array}{rcll}
F_X[x_2] & = & P([X \leq x_2] & \\
 & = & P[(X \leq x_1) \cup (x_1 < X \leq x_2)] & \text{disjuntos} \\
 & = & P[X \leq x_1] + P[x_1 < X < x_2] & \text{Axioma 3} \\
 & = & \underbrace{F_X[x_1]}_{\text{Definição}} + \underbrace{P[x_1 < X \leq x_2]}_{\text{número positivo}} &
\end{array}
$$

Logo

$$\boxed{F_x[x_2] \geq F_X[x_1]}$$

CDF

Propriedade 4

A CDF é uma função contínua à direita

$$\lim_{x \to x_o^+} F_X[x] = F_X[x_o]$$

Propriedade 5

A probabilidade de um intervalo é

$$P[a < X \leq b] = F_X[b] - F_X[a]$$

Prova: Supondo $a < b$, temos

$$\{-\infty < X \leq b\} = \{-\infty < X \leq a\} \cup \{a < X \leq b\}$$

Repare que os intervalos são disjuntos, ou seja,

$$P[-\infty < X \leq b] = P[-\infty < X \leq a] + P[a < X \leq b]$$

Logo

$$\boxed{P[a < X \leq b] = F_X[b] - F_X[a]}$$

Valor Esperado

Definição

Seja X uma variável discreta aleatória com os seguintes valores $a_1, \ldots, a_N$. O valor esperado ou média de X é o número $E(X)$ dado pela seguinte fórmula

$$E[X] = \sum_{i=1}^{N} a_i P[x = a_i]$$

Se os valores são igualmente válidos, ou seja, $P[x = a_i] = \frac{1}{N}$, então

$$E[X] = \frac{(a_1 + a_2 + \ldots + a_n)}{N}$$

Se a variável X toma infinitos valores então $E[X] = \sum_{i=1}^{\infty} a_i P[x = a_i]$. Aqui temos que nos preocupar se a série converge (Análise Matemática).

Variância

Definição

$$var[X] = \underbrace{E[X^2]}_{\sum_{i=1}^{N} a_i^2 P[x=a_i]} - E^2[X]$$

A variância mostra o quão espalhados estão os valores em torno da média.

Proposição 3.1

Seja X uma variável aleatória discreta com $E[X] = \mu$, então

$$Var[X] = E[(X-\mu)^2] = \sum_{i=1}^{N}(a_i - \mu)^2 P[X = a_i]$$

Prova:

$$\begin{aligned}
\sum_{i=1}^{N}(a_i-\mu)^2 P[X=a_i] &= \sum_{i=1}^{N}(a_i^2+\mu^2-2a_i\mu)P[X=a_i] \\
&= \underbrace{\sum_{i=1}^{N} a_i^2 P[x=a_i]}_{E[X^2]} - 2\mu \underbrace{\sum_{i=1}^{N} a_i P[x=a_i]}_{E[X]} + \\
&\quad \mu^2 \underbrace{\sum_{i=1}^{N} P[X=a_i]}_{1}
\end{aligned}$$

Mas $\mu = E[X]$, logo

$$\sum_{i=1}^{N}(a_i-\mu)^2 P[X=a_i] = E[X^2]-2E[X]E[X]+E[X]^2 = E[X^2]-E^2[X] \quad \square$$

Exemplo

Um moeda equilibrada é lançada 3 vezes. X é o número de caras. Qual é o valor esperado e a variância?

Solução:

$$\Omega = \{\underbrace{HHH}_{3}, \underbrace{HHT}_{2}, \underbrace{HTH}_{2}, \underbrace{HTT}_{1}, \underbrace{THH}_{2}, \underbrace{THT}_{1}, \underbrace{TTH}_{1}, \underbrace{TTT}_{0}\}$$

$$X = \begin{cases} 0 & \rightarrow & p = \frac{1}{8} \\ 1 & \rightarrow & p = \frac{3}{8} \\ 2 & \rightarrow & p = \frac{3}{8} \\ 3 & \rightarrow & p = \frac{1}{8} \end{cases}$$

O valor esperado é:

$$E[X] = 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8} = \frac{3}{2}$$

Exemplo (cont.)

A variância é:

$$Var[X] = 0^2 \times \frac{1}{8} + 1^2 \times \frac{3}{8} + 2^2 \times \frac{3}{8} + 3^2 \times \frac{1}{8} - \left(\frac{3}{2}\right)^2 = \frac{3}{4}$$

Variável Indicadora

Definição

Seja A um evento no espaço de probabilidades Ω. Com A nós associamos uma variável aleatória I_A (apenas uma função em Ω) tal que:

$$I_A(s) = \begin{cases} 1, & \text{se } s \in A \\ 0, & \text{se } s \notin A \end{cases}$$

onde s é um elemento em Ω.

A variável aleatória I_A é variável indicadora de A.

Exemplo

Se $x \sim Geo\left(\frac{1}{2}\right)$, determine $E[I_{(2,6)}(X)]$
Solução: Sabemos que para a Geométrica $p_X[k] = (1-p)^{k-1}p$ para $k = 1, 2, \ldots$. Como a função indicadora reduz o espaço de k para $k = 3, 4, 5$, temos:

$$\begin{aligned} E[I_{(2,6)}(X)] &= \sum_{k=3}^{5} 1 \times P(X = k) \\ &= \left((1-p)^2 + (1-p)^3 + (1-p)^4\right) p \end{aligned}$$

Para $p = \frac{1}{2}$, $E[X] = \frac{7}{32}$.

Valor Esperado e Variância

Bernoulli

$$X \sim Ber(p)$$

x	0	1
$P[X = x]$	q	p

mas $q = 1 - p$.
O valor esperado é dado por

$$E[X] = 0 \times q + 1 \times p = p$$

E a variância por

$$Var[X] = 0^2 \times q + 1^2 \times p - p^2 = p - p^2 = pq$$

Valor Esperado - Binomial

Lembrando que a variável aleatória binomial conta o número de sucessos em M tentativas independentes com a variável Bernoulli (p).

A variável aleatória $X \sim Bin(M, p)$ toma os valores $0, 1, 2, \ldots, M$. A PMF de X é dada por

$$P[X = k] = \binom{M}{k} q^{M-k} p^k$$

para $k = 0, 1, 2, \ldots, M$, onde $q = 1 - p$.
Por exemplo para $Bin(4, p)$, temos:

k	0	1	2	3	4
$P[X = k]$	q^4	$4q^3p$	$6q^2p^2$	$4qp^3$	p^4

Valor Esperado - Binomial (cont.)

Note também que se somarmos todas as probabilidades

$$\sum_{k=0}^{M} \binom{M}{k} q^{M-k} p^k = (q+p)^M = 1$$

Finalmente, se $X \sim Bin(M, p)$ então

$$\begin{aligned} E[X] &= Mp \\ Var[X] &= Mpq \end{aligned}$$

Prova:

1º Método: Suponha uma moeda com probabilidade p de sair cara. A mesma é lançada M vezes e o número de caras é contado.

Valor Esperado - Binomial (cont.)

$$X \sim Bin(M, p)$$

Seja X_k uma variável aleatória definida por

$$X_k = \begin{cases} 1 & \text{se for cara no k-ésimo lançamento} \\ 0 & \text{se for coroa no k-ésimo lançamento} \end{cases}$$

Podemos, então, escrever

$$X = \underbrace{X_1 + X_2 + \ldots + X_M}_{\text{Independentes } Ber(p)}$$

Sabemos que para cada um deles temos:

$$\begin{aligned} E[X_i] &= p \\ Var[X_i] &= pq \end{aligned}$$

Valor Esperado - Binomial (cont.)

Podemos escrever (e depois precisamos provar):

$$\begin{aligned} E[X] &= E[X_i] + \ldots + E[X_M] \\ Var[X] &= Var[X_i] + \ldots + Var[X_n] \end{aligned}$$

Logo:

$$\begin{aligned} E[X] &= Mp \\ Var[X] &= Mpq \end{aligned}$$

2º Método: Livro do Kay, página 138

Valor Esperado - Binomial (cont.)

$$
\begin{aligned}
E[X] &= \sum_{k=0}^{M} kP[X=k] \\
&= \sum_{k=0}^{M} k\binom{M}{k} p^k (1-p)^{M-k} \\
&= \sum_{k=0}^{M} k \frac{M!}{(M-k)!k(k-1)!} p^k (1-p)^{M-k} \\
&= \sum_{k=0}^{M} \frac{M!}{(M-k)!(k-1)!} p^k (1-p)^{M-k} \\
&= \sum_{k=0}^{M} \frac{M(M-1)!}{(M-k)!(k-1)!} pp^{k-1} (1-p)^{M-1-(k-1)}
\end{aligned}
$$

Valor Esperado - Binomial (cont.)

Logo

$$E[X] = Mp\sum_{k=0}^{M}\frac{(M-1)!}{(M-1-(k-1))!(k-1)!}p^{k-1}(1-p)^{M-1-(k-1)}$$

Fazendo $N = M-1$ e $l = k-1$, temos

$$\begin{aligned} E[X] &= Mp\underbrace{\sum_{\substack{l=-1\\ l=0}}^{N}\frac{N!}{(N-l)!l!}p^{l}(1-p)^{N-l}}_{1} \\ E[X] &= Mp \end{aligned}$$

$Var[X]$ - complicado! Vamos usar a função característica.

$$Hg(n, M, N)$$

Suponha que tenhamos N bolas numa caixas, das quais M são vermelhas (são amostradas m bolas sem reposição). Seja a variável aleatória X o número de bolas vermelhas na amostra. Tal X é chamado de variável aleatória hipergeométrica.

A variável X pode tomar qualquer dos valores em $0, 1, 2, 3 \ldots, n$.

$$P[X = k] = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$$

Sem provar, podemos escrever

Hipergeométrica (cont.)

$$
\begin{aligned}
E[X] &= n\left(\frac{M}{N}\right) \\
Var[X] &= n\left(\frac{M}{N}\right)\left(\frac{N-M}{N}\right)\left(\frac{N-n}{N-1}\right)
\end{aligned}
$$

Se $p = \frac{M}{N}$ com $M < N$ e consequentemente $q = 1 - p = \frac{N-M}{N}$, logo

$$
\begin{aligned}
E[X] &= np \\
Var[X] &= npq \quad \underbrace{\left(\frac{N-n}{N-1}\right)}_{\text{fator de correção}}
\end{aligned}
$$

Se n é pequeno comparado com N, então

Hipergeométrica (cont.)

$$\frac{N-n}{N-1} \approx 1 \rightarrow Bin(n, \tfrac{M}{N})$$

Geométrica

- Parecida com a Binomial.
- Suponha, novamente, que uma moeda é lançada e que a probabilidade de sair cara é p.
- Ao invés de um número fixo de lançamentos, o experimento acaba quando, por exemplo, cara aparece pela primeira vez. Contamos, então, o número de vezes que a moeda foi lançada. Os valores da variável são

$$1, 2, \ldots$$

  Pode ser que nunca consigamos uma cara, mas esta possibilidade é "infinitamente impossível".
- O número de tentativas de Bernoulli até o primero sucesso é uma variável aleatória geométrica.

$$P[X = k] = q^{k-1}p$$

Geométrica (cont.)

Verificando se a soma das probabilidades é 1

$$\sum_{k=1}^{\infty} q^{k-1}p = p + pq + pq^2 + \ldots = \frac{p}{1-q} = 1$$

Se $X \sim Geom(p)$, então

$$\begin{aligned} E[X] &= \frac{1}{p} \\ Var[X] &= \frac{q}{p^2} \end{aligned}$$

Prova para $E[X]$ - Kay, página 139

Geométrica (cont.)

$$
\begin{aligned}
E[X] &= \sum_{k=1}^{\infty} k q^{k-1} p \\
&= p \sum_{k=1}^{\infty} \underbrace{k q^{k-1}}_{\frac{dq^k}{dq}} \\
&= p \frac{d \sum_{k=1}^{\infty} q^k}{dq} \\
&= p \frac{d \frac{q}{1-q}}{dq} \\
&= p \frac{(1-q) - q(-1)}{(1-q)^2} = p \frac{1}{(1-q)^2} \\
&= \frac{1}{p} \quad \square
\end{aligned}
$$

Variáveis Aleatórias

Poisson

$$X \sim Pois(\lambda)$$

Para Poisson, o valor esperado e a variância são:

$$\begin{aligned} E[X] &= \lambda \\ Var[X] &= \lambda \end{aligned}$$

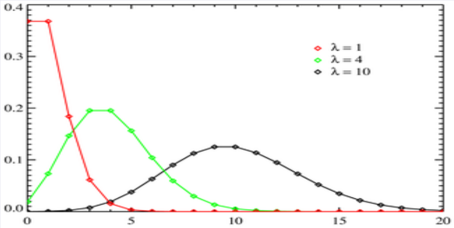


Figura 15: PMF Poisson

Valor Esperado

Propriedades do Valor Esperado

Nem todas as PMFs com uma infinidade de valores possuem valor esperado. Para ter é preciso que:

a) Soma - absolutamente somável
b) Soma dos valores absolutos dos termos é finito.

$$E[|X|] = \sum_{k=-\infty}^{k=\infty} |k| p_X[k] < \infty$$

Propriedades do Valor Esperado

- PMF simétrica - valor esperado localizado no centro da PMF.
- O valor esperado geralmente não indica o valor mais provável
- Mais de uma PMF apresenta o mesmo valor esperado

Exemplo de Valor Esperado não-finito

Considere a variável aleatória com distribuição **zipf** usada para a análise da popularidade de web sites e web caching. Suponha que $P(X = k) = \frac{C^{-1}}{k^2}$ com $k = 1, 2, \ldots,$ onde

$$C = \sum_{n=1}^{\infty} \frac{1}{n^2}$$

Calcular o valor esperado $E[X]$.

Solução: Vamos aproveitar e verificar se $P(X = k)$ é mesmo uma PMF. Para isto, devemos lembrar que:

- $\sum_{n=1}^{\infty} \frac{1}{n^2}$ nada mais é do que uma série do tipo $\sum_{n=1}^{\infty} \frac{1}{n^r}$ onde para $r > 1$ a série converge.
- O resultado da soma da série é a função zeta de Riemann.
- $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$ cuja a prova foi dada por Euler (Basel problem).

Exemplo de Valor Esperado não-finito (cont.)

Para mostrar que a soma de $P(X = k)$ para $k = 1, 2, \ldots$ é 1, vamos usar o resultado de Euler duas vezes:

$$\begin{aligned}
\sum_{k=1}^{\infty} \frac{C^{-1}}{k^2} &= C^{-1} \sum_{k=1}^{\infty} \frac{1}{k^2}, \text{ lembrando que } C \text{ é uma constante} \\
&= \frac{6}{\pi^2} \sum_{k=1}^{\infty} \frac{1}{k^2}, \text{ o somatório é a mesma série} \\
&= \frac{6}{\pi^2} \times \frac{\pi^2}{6} \\
&= 1
\end{aligned}$$

No caso do valor esperado, temos:

Exemplo de Valor Esperado não-finito (cont.)

$$
\begin{aligned}
E[X] &= \sum_{k=1}^{\infty} k \times P(X = k) \text{ (definição)} \\
&= \sum_{k=1}^{\infty} k \frac{C^{-1}}{k^2} \\
&= C^{-1} \sum_{k=1}^{\infty} k \frac{1}{k^2} \\
&= C^{-1} \sum_{k=1}^{\infty} \frac{1}{k}, \text{ série geométrica - diverge}
\end{aligned}
$$

Como consequência, $E[X^r]$ para $r \geq 1$ é infinito.

Outro Exemplo de Valor Esperado

Considere neste exemplo o mesmo C do caso anterior, mas

$$P(X = k) = P(X = -k) = \frac{1}{2} \times \frac{C^{-1}}{k^2} \text{ para } k = 1, 2, \ldots,$$

claramente $\sum_{k=1}^{\infty} P(X = k)$ é 1.
No caso do valor esperado, temos

$$\begin{aligned}
E[X] &= \sum_{k=1}^{\infty} kP(X = k) + \sum_{k=-\infty}^{-1} kP(X = k) \\
&= \frac{1}{2C} \sum_{k=1}^{\infty} \frac{1}{k} + \frac{1}{2C} \sum_{k=-\infty}^{-1} \frac{1}{k} \\
&= \frac{\infty}{2C} + \frac{-\infty}{2C} \\
&= \text{indefinido ou zero?}
\end{aligned}$$

Considere

$$Y = g(X)$$

então

$$\begin{array}{lcll} E[Y] & = & \sum_i y_i p_Y[y_i] & \text{Definição} \\ E[g(X)] & = & \sum_i g(x_i) p_X[x_i] & \end{array}$$

Exemplo 1

Dado

$$g(X) = aX + b$$

deseja-se saber $E[g(X)] = E[aX + b]$

Solução: aplicando a definição

Valor Esperado para uma função de uma variável aleatória (cont.)

$$
\begin{aligned}
E[g(X)] &= \sum_i (ax_i + b)\, p_X[x_i] \\
&= a \sum_i x_i p_X[x_i] + b \underbrace{\sum_i p_X[x_i]}_{1} \\
&= aE[X] + b
\end{aligned}
$$

Exemplo 2 Dado

$$g(X) = \alpha g_1(X) + \beta g_2(X)$$

deseja-se saber $E[g(X)] = E[\alpha g_1(X) + \beta g_2(X)]$

Solução: Aplicando a definição

Valor Esperado para uma função de uma variável aleatória (cont.)

$$\begin{aligned} E[g(X)] &= E[\alpha g_1(X) + \beta g_2(X)] \\ &= \alpha \sum_i g_1(x_i) p_X[x_i] + \beta \sum_i g_2(x_i) p_X[x_i] \\ &= \sum_i \left[\alpha g_1(x_i) + \beta g_2(x_i)\right] p_X[x_i] \\ &= \sum_i g(x_i) p_X[x_i] \\ &= E[g(x)] \end{aligned}$$

Conclusão: O valor esperado é um operador linear.

Exemplo 2

Dado

$$p_X[k] = \frac{1}{5} \quad k = 0, 1, 2, 3, 4$$

achar $E[Y]$ onde $Y = g(X) = \sqrt{X}$

Valor Esperado para uma função de uma variável aleatória (cont.)

Solução: Usando a definição

$$
\begin{aligned}
E[\sqrt{X}] &= \textstyle\sum_{k=0}^{4} \sqrt{k} p_X[k] \\
&= \textstyle\sum_{k=0}^{4} \sqrt{k} \times \frac{1}{5} \\
&= \textstyle\frac{1}{5} \sum_{k=0}^{4} \sqrt{k} \\
&= \frac{3+\sqrt{2}+\sqrt{3}}{5}
\end{aligned}
$$

Considere o seguinte

$$
\begin{aligned}
E[X] &= \textstyle\sum_{k=0}^{4} k p_X[k] \\
&= \textstyle\sum_{k=0}^{4} k \frac{1}{5} \\
&= \textstyle\frac{1}{5} \sum_{k=0}^{4} k = \frac{10}{5} = 2
\end{aligned}
$$

Usando o resultado acima, podemos escrever

$$\sqrt{E[X]} = \sqrt{2}$$

Valor Esperado para uma função de uma variável aleatória (cont.)

Repare que

$$\sqrt{E[X]} \neq E[\sqrt{X}]$$

ou seja, o valor esperado não é um operador comutativo.

Exemplo 4

Considere

$$\begin{cases} b & \rightarrow & \text{valor a ser predito} \\ X & \rightarrow & \text{variável aleatória} \end{cases}$$

Objetivo: Achar um valor que, em média, seja próximo do verdadeiro valor da variável X.

Solução

Valor Esperado para uma função de uma variável aleatória (cont.)

- Vamos definir a função erro como

$$\text{erro} = X - b$$

Uma boa medida é

$$(X - b)^2$$

- O que queremos é que

$$E[(X - b)^2] \quad \text{ou seja, o MSE (Mean Square Error)}$$

seja o menor possível.

Valor Esperado para uma função de uma variável aleatória (cont.)

- Calculando o MSE

$$
\begin{aligned}
MSE(b) &= E[(X-b)^2] \\
&= E[X^2 - 2Xb + b^2] \\
&= E[X^2] - 2E[Xb] + E[b^2] \\
&= E[X^2] - 2bE[X] + b^2
\end{aligned}
$$

- Precisamos achar o valor que minimiza o MSE

$$
\frac{dMSE(b)}{db} = -2E[X] + 2b = 0
$$

$$
\downarrow
$$

$$
E[X] = b
$$

Valor Esperado para uma função de uma variável aleatória (cont.)

logo

$$
\begin{aligned}
MSE(b) &= E[(X-b)^2] \\
&= E[(X-E[X])^2] \\
&= Var[X]
\end{aligned}
$$

Propriedades da Variância

Seja c uma constante, então:

- $Var[c] = 0$
- $Var[X + c] = Var[X]$
- $Var[cX] = c^2 Var[X]$

$E[X]$ e $E[X^2]$ são o primeiro e segundo momentos, respectivamente.

$$
\begin{array}{lcl}
E[X^n] & \rightarrow & \text{n-ésimo momento} \\
E[X^n - E^n[X]] & \rightarrow & \text{momentos centrais}
\end{array}
$$

A variância é um operador não-linear.

Definição: A função característica da variável X é definida como

$$\phi_X(\omega) = E[e^{\jmath\omega X}]$$

Repare que o argumento de $E[\bullet]$ é uma transformação de uma variável aleatória, ou seja,

$$g(X) = e^{\jmath\omega X}$$

Usando a definição podemos escrever

$$\begin{aligned} \phi_X(\omega) &= E[e^{\jmath\omega X}] \\ &= \sum_i e^{\jmath\omega x_i} p_X[x_i] \\ &= \sum_i p_X[x_i] e^{\jmath\omega x_i} \end{aligned}$$

repare que nada mais é do que a Transformada Discreta de Fourier

Função Característica (cont.)

A idéia é usar a função característica para calcular $E[X]$. Para isso considere

$$\begin{aligned}
\frac{d\phi_X(\omega)}{d\omega} &= \frac{d}{d\omega}\sum_{k=-\infty}^{\infty} p_X[k]e^{\jmath\omega k} \\
&= \sum_{k=-\infty}^{\infty} p_X[k]\frac{de^{\jmath\omega k}}{d\omega} \\
&= \sum_{k=-\infty}^{\infty} p_X[k]\jmath k e^{\jmath\omega k}
\end{aligned}$$

Podemos substituir um valor determinado. Para os nossos propósitos, vamos usar $\omega = 0$

$$\begin{aligned}
\frac{1}{\jmath}\left.\frac{d\phi_X(\omega)}{d\omega}\right|_{\omega=0} &= \sum_{k=-\infty}^{\infty} k p_X[k]\underbrace{e^{\jmath\omega k}}_{1} \\
&= E[X]
\end{aligned}$$

Estendendo o resultado acima, podemos escrever

$$E[X^n] = \frac{1}{\jmath^n}\left.\frac{d^n\phi_X(\omega)}{d\omega^n}\right|_{\omega=0}$$

Variância - Geométrica

Lembrando que a PMF da VA Geométrica é dada por:

$$p_X[k] = (1-p)^{k-1}p$$

é possível determinar a função caraterística para a mesma. Para isso, considere

$$\begin{aligned}
\phi_X(\omega) &= \sum_{k=1}^{\infty} p_X[k] e^{\jmath\omega k} \\
&= \sum_{k=1}^{\infty} (1-p)^{k-1} p e^{\jmath\omega k} \\
&= p \sum_{k=1}^{\infty} (1-p)^{k-1} e^{\jmath\omega(k-1+1)} \\
&= p e^{\jmath\omega} \sum_{k=1}^{\infty} \left[(1-p)e^{\jmath\omega}\right]^{k-1} \\
&= p e^{\jmath\omega} \sum_{k'=0}^{\infty} \left[(1-p)e^{\jmath\omega}\right]^{k'} \quad \text{Isso é uma PG} \\
&= \frac{p e^{\jmath\omega}}{1-(1-p)e^{\jmath\omega}}
\end{aligned}$$

Variância - Geométrica (cont.)

Dado que a função característica foi determinada, podemos calcular os momentos e com eles a variância.

$$
\begin{aligned}
E[X^2] &= \frac{1}{\jmath^2} \left. \frac{d^2\phi_X(\omega)}{d\omega^2} \right|_{\omega=0} \\
&= \frac{2}{p^2} - \frac{1}{p}
\end{aligned}
$$

Logo

$$
\begin{aligned}
Var[X] &= E[X^2] - E^2[X] \\
&= \frac{2}{p^2} - \frac{1}{p} - \frac{1}{p^2} \\
&= \frac{1-p}{p^2} = \frac{q}{p^2}
\end{aligned}
$$

- A função característica sempre existe pois $|\phi_X(\omega)| < \infty$

  Prova:

Propriedades da Função Característica (cont.)

$$
\begin{aligned}
|\phi_X(\omega)| &= \left| \sum_{k=-\infty}^{\infty} p_X[k] e^{\jmath\omega k} \right| \\
&\leq \sum_{k=-\infty}^{\infty} \left| p_X[k] e^{\jmath\omega k} \right| \\
&\leq \sum_{k=-\infty}^{\infty} |p_X[k]| \underbrace{\left| e^{\jmath\omega k} \right|}_{1} \\
&\leq \sum_{k=-\infty}^{\infty} |p_X[k]| \\
&\leq 1
\end{aligned}
$$

Propriedades da Função Característica (cont.)

- A PMF pode ser recuperada da função característica

$$p_X[k] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_X(\omega) e^{-\jmath\omega k} d\omega$$

isso nada mais é do que a Transformada de Fourier Inversa da função característica.

Sabemos que a função característica está relacionada com a Transformada Discreta de Fourier, logo porque não pensar em Transformada $\mathcal{Z}$?

$$\phi_X(z) = E[z^X]$$

Exemplo: Função característica de X com distribuição de Poisson com parâmetro λ.

Solução

Uma Outra Maneira de ver a Função Característica (cont.)

$$\begin{aligned}
\phi_X(z) &= E[z^X] \\
&= \sum_{k=0}^{\infty} z^k P(X = k) \\
&= \sum_{k=0}^{\infty} z^k \frac{\lambda^k e^{-\lambda}}{k!} \\
&= e^{-\lambda} \sum_{k=0}^{\infty} \frac{(z\lambda)^k}{k!} \\
&= e^{-\lambda} e^{z\lambda} \\
&= e^{\lambda(z-1)}
\end{aligned}$$

Exemplo: Função característica de uma soma de variáveis aleatórias independentes.

Uma Outra Maneira de ver a Função Característica (cont.)

Solução

$$
\begin{aligned}
\phi_{X_1+\ldots+X_n}(z) &= E[z^{X_1+X_2+\ldots+X_n}] \\
&= E[z^{X_1}\ldots z^{X_n}] \\
&= E[z^{X_1}]\ldots E[z^{X_n}] \qquad \text{usando independência} \\
&= \phi_{X_1}(z)\ldots\phi_{X_n}(z)
\end{aligned}
$$

Exemplo: Utilização da função característica para o cálculo de probabilidades.

Solução: Considere a seguinte definição da função característica em termos da Transformada $\mathcal{Z}$

$$\phi_X(z) = E[z^X] = \sum_{k=0}^{\infty} z^k P(X = k)$$

Expandindo a série temos:

$$\phi_X(z) = P(X=0) + zP(X=1) + z^2P(X=2) + \dots$$

Quando $z = 0$

$$\phi_X(0) = P(X=0)$$

Derivando $\phi_X(z)$ uma vez, temos

$$\phi'_X(z) = P(X=1) + 2zP(X=2+\dots$$

Neste caso, quando $z = 0$, a função característica é

$$\phi'_X(0) = P(X=1)$$

Estendendo para derivadas de ordem superior, temos

$$\frac{\left.\phi_X^{(n)}(z)\right|_{z=0}}{n!} = P(X = n)$$

Exemplo: Se $\phi_X(z) = \left(\frac{1+z+z^2}{3}\right)^2$, encontre $P(X = 2)$.

Solução: Derivando duas vezes, temos

$$\begin{aligned}
\phi_X'(z) &= 2\left(\frac{1+z+z^2}{3}\right)\left(\frac{1+2z}{3}\right) \\
\phi_X''(z) &= 2\left(\frac{1+z+z^2}{3}\right)\left(\frac{2}{3}\right) + 2\left(\frac{1+2z}{3}\right)\left(\frac{1+2z}{3}\right)
\end{aligned}$$

Usando $\phi_X''(z)$ podemos facilmente achar que

Uma Outra Maneira de ver a Função Característica (cont.)

$$P(X=2)=\frac{\phi''_X(0)}{2!}=\frac{1}{2!}\left(\frac{4}{9}+\frac{2}{9}\right)=\frac{1}{3}$$