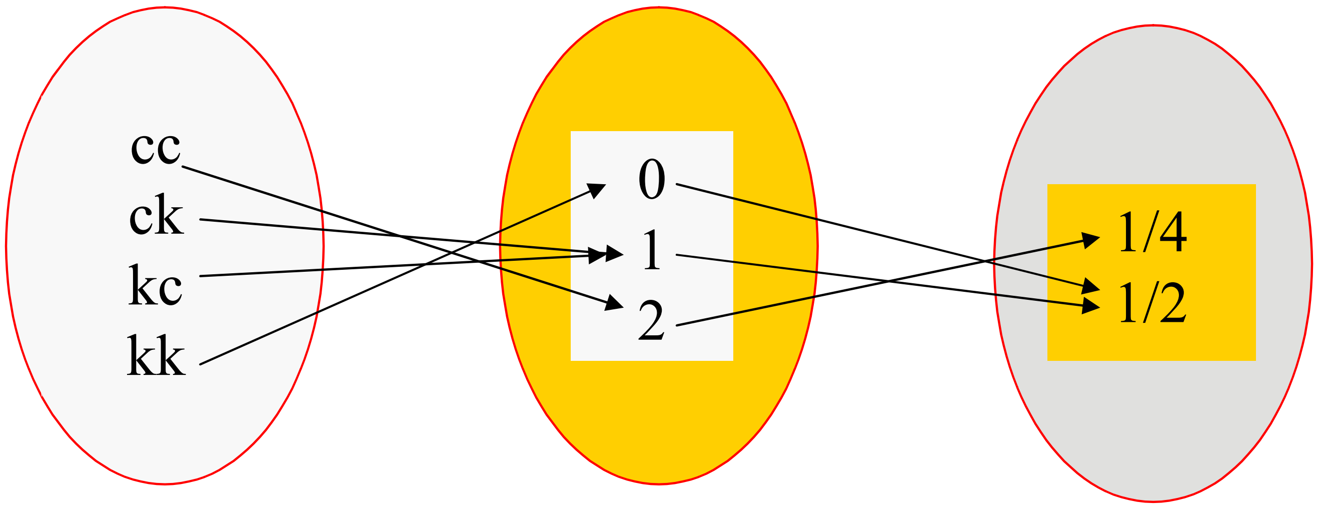
Variáveis Aleatórias - VA



Prof. Adriano Mendonça Souza, Dr.
Departamento de Estatística
- PPGEMQ / PPGEP - UFSM -

Introdução

Se entende por **VA** ou **V. indicadoras** uma lista de valores ou uma função genérica que associa um número conveniente aos eventos componentes do **S**, sejam qualitativos ou quantitativos.

Se $\boldsymbol{X = f(x)}$; a probabilidade a elas associada é indicada por $\boldsymbol{P[X = f(x)]}$.

$\boldsymbol{P(x = a)}$ indicará a probabilidade da **VA X** se encontrar num ponto qualquer entre **a** e **b**.

Exemplo

Seja o experimento “lançamento de 3 moedas”, determine a probabilidade de dar “pelo menos uma coroa”.

O espaço amostral desse experimento, visto abaixo, mostra os eventos possíveis e as respectivas freqüências ou probabilidades, agrupados pelo número de vezes que ocorre o evento “cara”.

X = VA = Número de caras

X	Número de caras	Eventos característicos	Freqüência
0	Nenhuma	KKK	1/8
1	Uma	CKK,KCK,KKC	3/8
2	Duas	KCC,CKC,CCK	3/8
3	Três	CCC	1/8
	C=cara K=coroa	(espaço amostral)	8/8

Definição 1

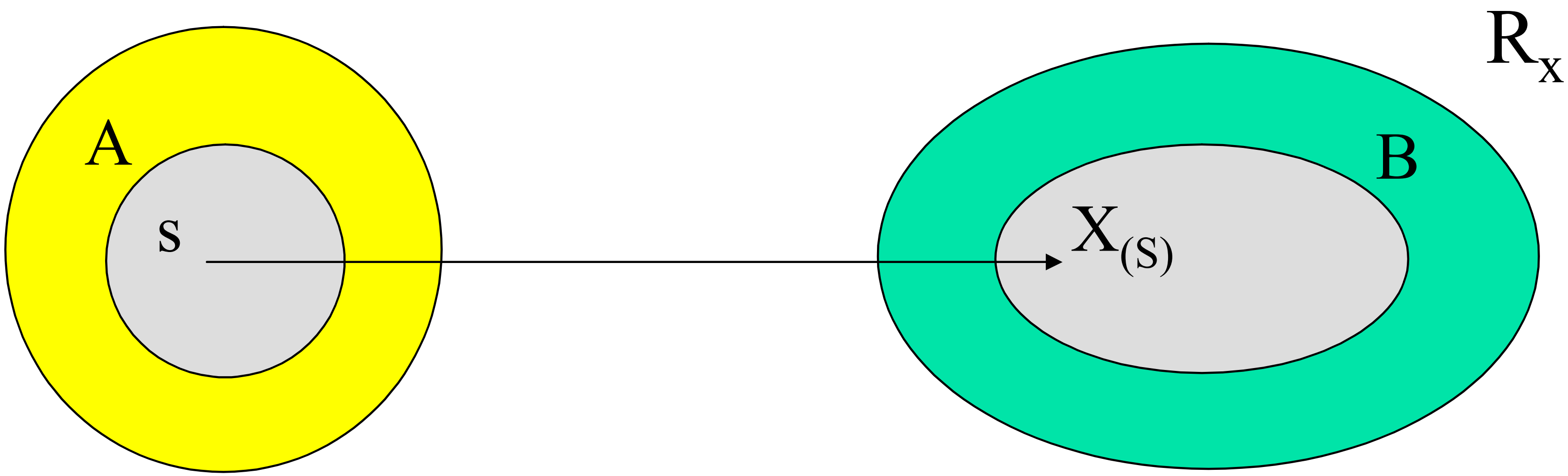
Seja um experimento ***E*** seu espaço amostral **S** e **X** uma **VA** definida em **S**, onde **Rx** é seu contradomínio.

Se **B** é um evento definido em relação a **Rx**, isto é, **B** $\subset$ **Rx**. Então, **A** será definido assim:

$$\mathbf{A} = \{ \mathbf{s} \in \mathbf{S} \;/\; \mathbf{X}(\mathbf{s}) \in \mathbf{B} \}$$

Explicando:

A será constituído por todos os resultados em **S**, para os quais **X(s)** $\in$ **B**. Neste caso, diremos que **A** e **B** são eventos equivalentes.



A e **B** serão equivalentes sempre que ocorrerem juntos, isto é, **A ocorreu, B ocorre** e vice-versa.

(A ocorre se e somente se X(s) ocorreu)

Exemplo

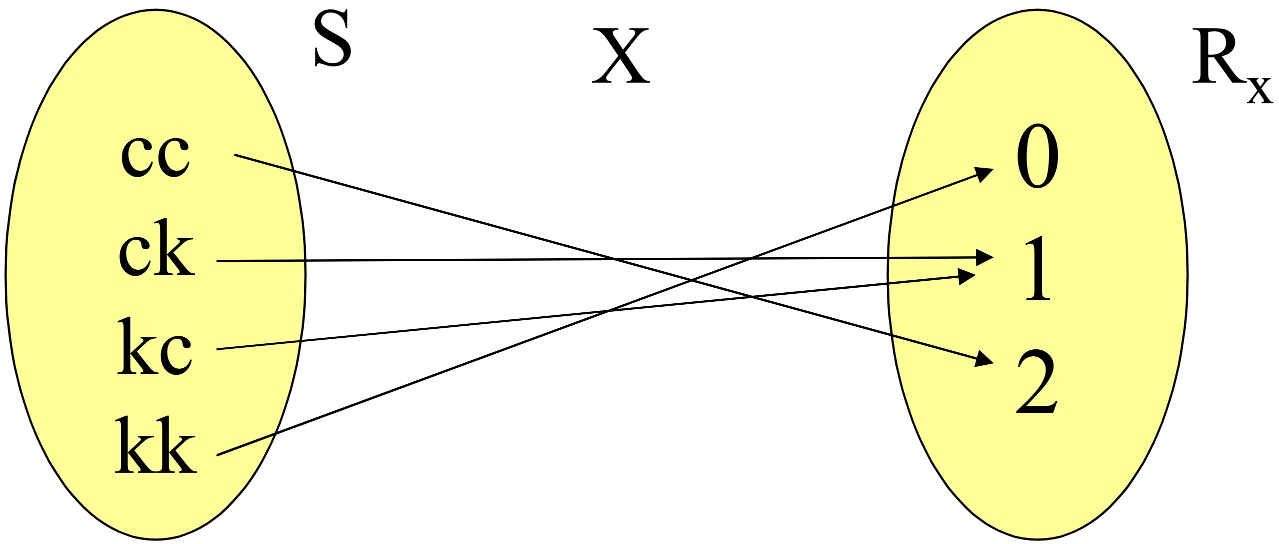
Seja o seguinte experimento:

E: Uma jogada de duas moedas

S = {cc, ck, kc, kk }

X: V.A: número de caras obtidas; Rx = {0, 1, 2 }

Seja B = {1}, logo X(kc) = 1 se e somente se X(s) = 1,
temos que A = {kc, ck} é equivalente a B.



Definição 2

Seja **B** um evento no contradomínio **Rx**.

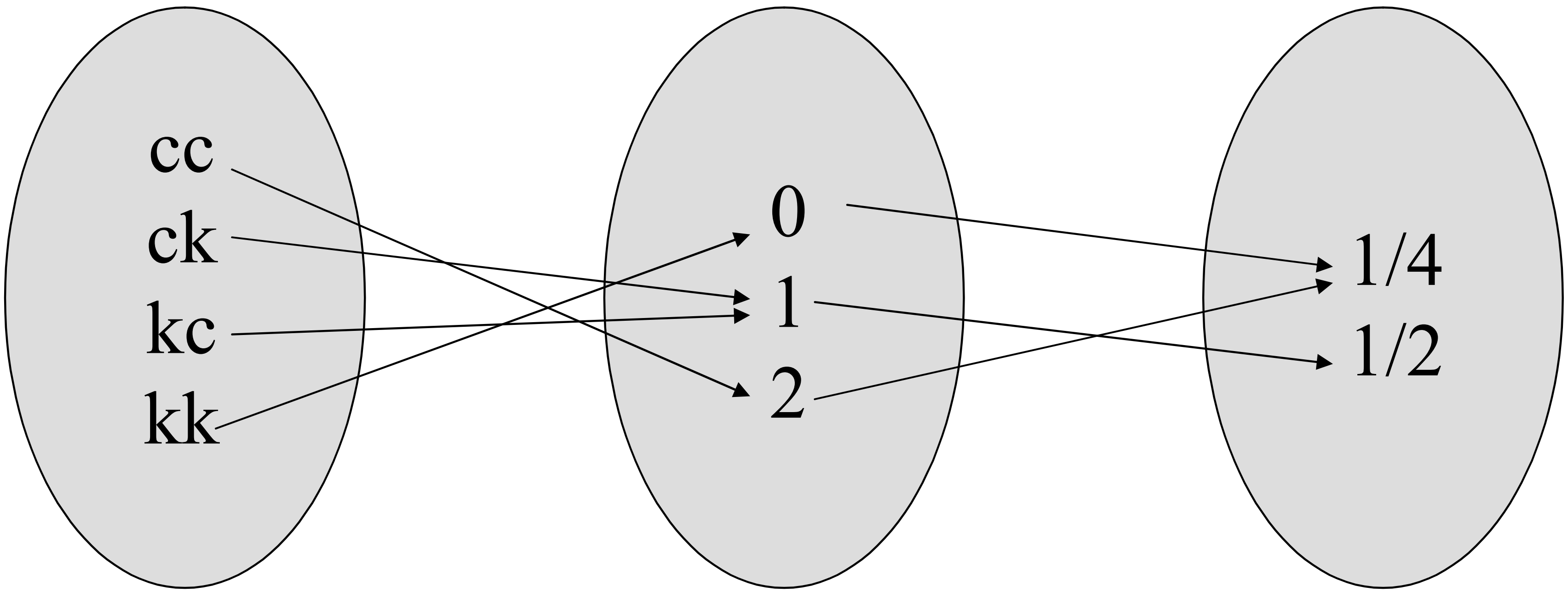
Nesse caso a **P(B)** será:

$$\mathbf{P(B) = P(A), \text{ onde } A = \{ s \in S / \ X(s) \in B \}}$$

Explicando: Definimos **P(B)** igual a probabilidade do evento $\mathbf{A} \subset \mathbf{S}$, o qual é equivalente a **B**, estamos admitindo que probabilidades possam ser associadas a eventos em **S**.

Portanto a definição acima torna possível atribuir probabilidades a eventos associados a **Rx** em termos de probabilidades definidas sobre **S**.

Então teremos:



Classificação das Variáveis Aleatórias

Variáveis Aleatórias Discretas - VAD

Seja **X** uma **VA**. Se o número de valores possíveis de **X** (isto é, Rx, o contradomínio) for finito ou infinito numerável, denominamos **X** de **VAD**.

Isto é, os valores possíveis de X podem ser postos em listas como $x_1, x_2, \ldots, x_n$.

Vide os exemplos anteriores.

Distribuição de probabilidade de uma VAD

Uma distribuição de probabilidade é uma distribuição de freqüência relativa ou probabilidades para os resultados do espaço amostral.

Mostra a proporção de vezes em que a VA tende a assumir cada um dos diversos valores.

Definição

Seja **X** uma **VAD** a cada resultado x_i possível do contradomínio pudermos associar um número

$P(x_i) = P(X = x_i)$, denominamos de probabilidade de x_i.

Os números ***p(xi)***, i = 1, 2,..., n devem satisfazer:

a) $P(x_i) \geq 0$

b) $\sum_{i=1}^{\infty} P(x_i) = 1$

A função ***P***, definida acima, é denominada de função de probabilidade ou função de probabilidade no ponto da VA X.

A coleção de pares [*xi*, *P(xi)*], *i = 1, 2,..., n*
é denominada de distribuição de probabilidade.

Exemplo

Considere a seguinte variável aleatória:

X: VAD: número de caras em duas jogadas de uma moeda

Resultado (S)	Valor (VAD)
cc ck kc kk	2 1 0 0

Sendo a moeda honesta *p(c)* = *p(k)* = *1/2*, logo:

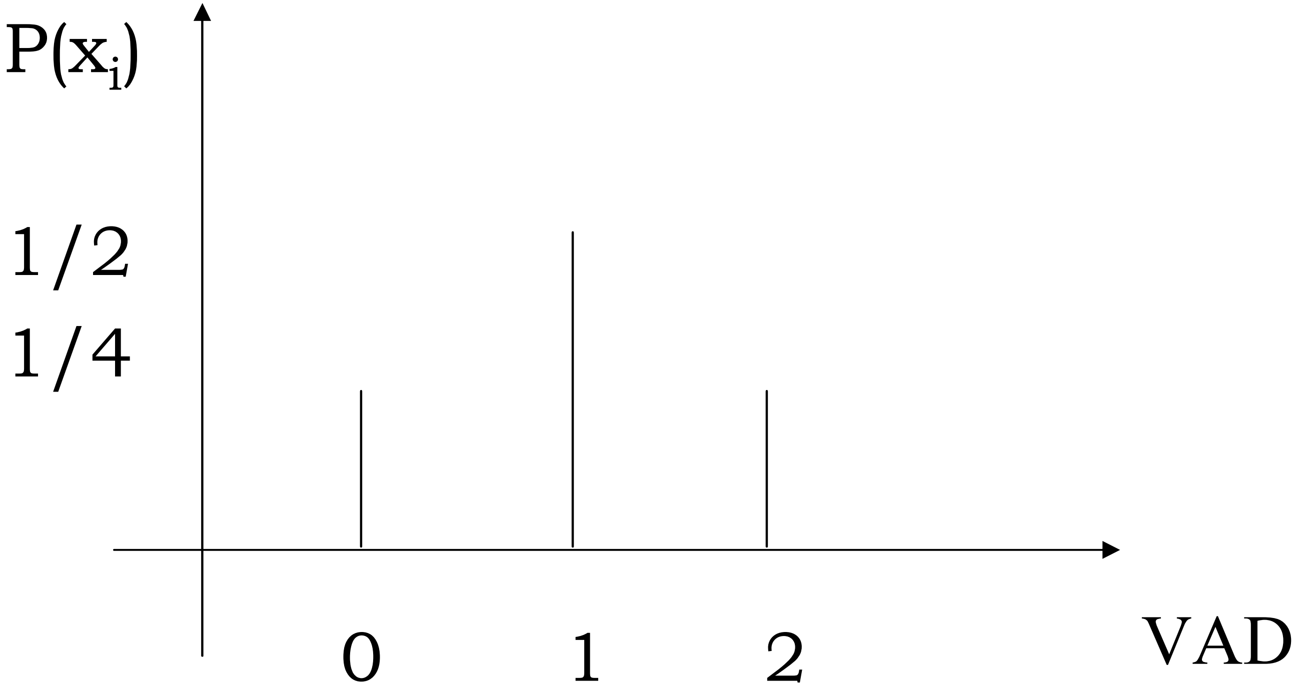
Resultado(S)	Probab. do resultado	Num. de caras(VAD)	P(x)
cc	1/2. 1/2 = ¼	2	0,25
ck	1/2. 1/2 = ¼	1	0,50
kc	1/2. 1/2 = ¼		
kk	1/2. 1/2 = ¼	0	0,25

logo a distribuição de probabilidade será:

VAD (x)	0	1	2
$P(x_i)$	0,25	0,50	0,25

$\Rightarrow p(xi) \geq 0$ $\Rightarrow \sum p(xi) = 1{,}00$

Graficamente a distrib. de probabilidade fica:



Função de distribuição ou Função acumulada de distribuição F(X) de uma VAD

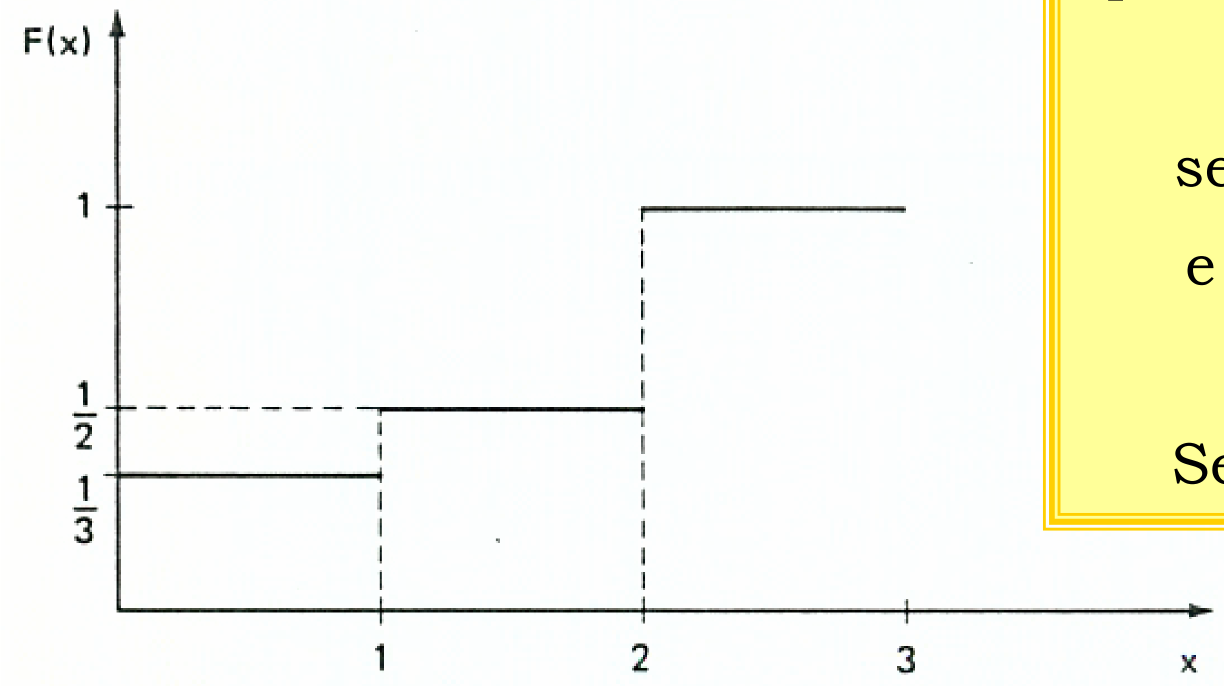
Definição

Seja **X** uma **VAD**, defini-se a função ***F(X)*** como a função distribuição acumulada da **VA X** (abreviadamente ***fd***), a ***F(x)*** de uma **VA X** é definida como probabilidade de **X** ser menor ou igual a um determinado o valor de **x**, i. é., $F(X) = P(X \leq x)$

Teorema

Se **X** for **VAD**, ***F(X)*** = $\sum_j p(x_j)$, onde o somatório é estendido a todos os índices j que satisfaçam a condição de $X_j \leq x$.

Graficamente temos F(X):



A função F(x)

possui segmentos de retas

horizontais

se tivermos uma **VAD**

e segmentos de retas

contínuos

Se tivermos uma **VAC**

Exemplo

No lançamento de dois dados, onde a soma de pontos que aparecem para cima é a VAD considerada, a sua F(X), será:

$F(2) = P(x \leq 2) = P(x = 2) = 1/36$

$F(3) = P(x \leq 3) = P(x = 2) = + P(x = 3) = 1/36 + 2/36 = 3/36$

$F(4) = P(x \leq 4) = P(x = 2) + P(x = 3) + P(x = 4) = 1/36 + 2/36 +$
$+ 3/36 = 6/36$

...

...

...

$F(12) = P(x + 2) + P(x = 3) + \ldots + P(x = 12) = 36/36 = 1$

Como ficaria o gráfico desta distribuição?

Variáveis Aleatórias Contínuas - VAC

Distribuição de Probabilidade Contínua

É aquela que pode tomar qualquer valor em um determinado intervalo. Diz-se que **X** é uma **VAC**, se existir uma função ***f(x)***, denominada função densidade de probabilidade (***fdp***) de **x** que satisfaça às seguintes condições:

a) $f(x) \geq 0$ para todo o x;

b) $\int_{-\infty}^{+\infty} f(x)\,dx = 1$;

c) para quaisquer a, b, com $-\infty < a < b +\infty$, teremos que

$P(a \leq x \leq b) = \int_a^b f(x)\,dx$ e para $a < b$ em Rx, teremos que

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

Observações:

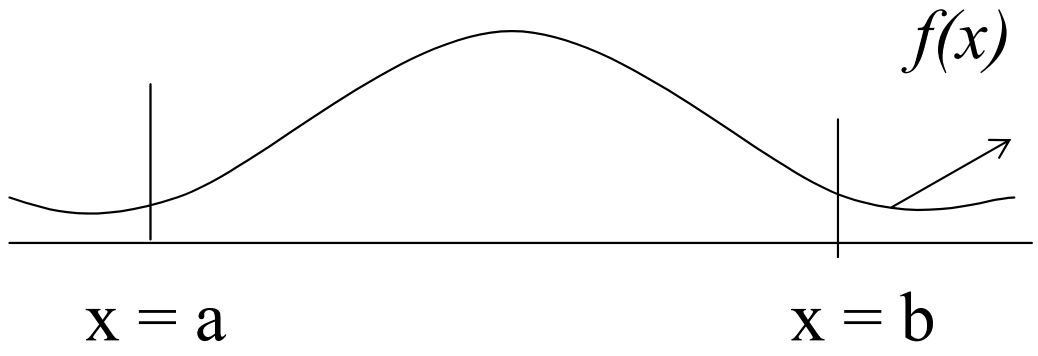
1) A definição acima mostra que qualquer valor específico de **X**, ***P(X = Xo) = 0***, pois:

$$P(X = Xo) = \int_{x_0}^{x_0} f(x)dx = 0$$

Sendo assim, as probabilidades abaixo serão todas iguais se **X** for uma **VAC**.

$$P(a \le x \le b),\ P(a \le x < b),\ P(a < x \le b),\ P(a < x < b)$$

2) ***f(x)***, densidade de probabilidade, não é probabilidade, somente será probabilidade se integrada entre dois limites, **produzindo uma área sob função**.



Função de distribuição ou Função acumulada de distribuição F(X) de uma VAC

Quanto à função de Repartição em uma **VAC**, neste caso ela é definida como:

$$F(x): \int_{-\infty}^{x} f(x)\, dx$$

Pode-se provar que f(x) = $\dfrac{d\, f(x)}{dx}$ *para todo **x** no qual **F** seja derivável. Como **F(+** ∞**) = 1**, devemos ter sempre* $\int_{-\infty}^{+\infty} f(x)\, dx = 1$*, isto é, a **área total debaixo da curva de probabilidade vale sempre 1**.*

Exemplo

Suponhamos que a **VA X** tome os valores **0, 1, 2** com probabilidades **1/3, 1/6, 1/2** respectivamente, encontre a sua função repartição **F(X)**.

Então:

$F(X) = 0 \quad$ se $X \angle 0$

$F(X) = 1/3 \quad$ se $0 \leq X \angle 1$

$F(X) = 1/2 \quad$ se $1 \leq X \angle 2$

$F(X) = 1 \quad$ se $X \geq 2$

Propriedades da Função Acumulada ou Repartição F(X)

Sendo $F(X) = p(x_i \leq x)$ temos as seguintes propriedades:

1. $F(X) = \sum p(x_i)$

2. $F(-\infty) = 0$

3. $F(+\infty) = 1$

4. $P(a \langle x \leq b) = F(b) - F(a)$

5. $P(a \leq x \leq b) = F(b) - F(a) + P(x = a)$

6. $P(a \langle x \langle b) = F(b) - F(a) - P(x = b)$

7. $F(x)$ é uma contínua a direita, logo $F(x)$ é descontínua à esquerda, nos pontos em que a probabilidade é diferente de zero para $P(X = Xo) \neq 0$

8. A função é não decrescente, isto é, $F(b) \geq F(a)$, pois:

$$P(a \langle x \leq b) = F(b) - F(a) \geq 0$$

Variáveis Absolutas e Acumuladas

As **VA** podem ser associadas tanto a valores absolutos quanto a valores acumulados.

Em valores **absolutos**, sua indicação será $\boldsymbol{X = f(x)}$;

Em valores **acumulados**, sua indicação será $\boldsymbol{X < f(x)}$;

O comportamento probabilístico corresponde a cada uma das situações é diferente e as distribuições de probabilidade levarão em conta esse fato.

Grandezas Características

As **VA** podem ser caracterizadas pelas grandezas:

Média
Moda
Mediana
Dispersão

A medida de **tendência central** é uma média ponderada que recebe o nome particular de **esperança matemática** (ou valor esperado);

as grandezas de dispersão são a variância e o desvio- padrão.

Esperança Matemática (E(x))

Seja uma VAD X toma os valores x_1, x_2, ..., x_n com as probabilidades correspondentes p_1, p_2,..., p_n, então:

$$m = E(X) = x_1 \, . \, p_1 + x_2 \, . \, p_2 + \ldots + x_n \, . \, p_n = \sum X_n . P_n$$

Seja uma VAC definida por $X = f(x)$, a expressão resultante é obtida trocando-se a operação “soma das variáveis” pela integral da função”, então:

$$m = E(x) = \mu x = x = \int_{-\infty}^{+\infty} x \, . \, f(x) \, dx$$

Em uma analogia com as grandezas da física, associa-se a **esperança matemática** a uma espécie de **centro de gravidade probabilístico**, em torno do qual se distribuem os valores da **VA**.

Uma das aplicações mais conhecidas da esperança matemática é a determinação do valor das apostas em jogos de azar, em termos de seguro, é utilizada no cálculo do desembolso esperando pelo segurador nas suas operações.

Exemplo

E: Um dado equilibrado é lançado

$S = \{1, 2, 3, 4, 5, 6,\}$ e $P(x_i) = 1/6$

X: VAD: Número de pontos obtidos

$E(x) = 1.1/6 + 2.1/6 + 3.1/6 + 4.1/6 + 5.1/6 + 6.1/6 = 3{,}5$

- ***E(x)*** não é o resultado que podemos esperar quando **x** for observado uma única vez.
- ***E(x)*** nem mesmo é um valor possível de x;

Se obtivermos um grande número de observações independentes de X (VAD X) e calcula-se a média aritmética desses resultados sob condições bastante gerais, a média aritmética será próxima de E(x) em sentido probabilístico.

Propriedades da Expectância

1) Se x = c, onde "c" é uma constante, então:

$$\mathbf{E(x) = c}$$

Seja ela uma VAC ou uma VAD

$$E(x) = \int_a^b c \,.\, dx = c \int_a^b dx = c\,[x]_a^b$$

2) Se "c" for uma constante e "x" uma VAC, então:

$$E(cx) = c\,E(x)$$

$$E(cx) = \int_a^b c \,.\, x \,.\, f(x)\, dx = c \int_a^b x \,.\, f(x) \,.\, dx = c \,.\, E(x)$$

3) Sejam X e Y duas V.A. quaisquer, X e Y não associadas, então:

$$E(x+y) = E(x) + E(y)$$

$$E(x+y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x+y).f(x,y)dxdy = \int_{-\infty}^{+\infty}x.f(x,y)dx + \int_{-\infty}^{+\infty}y.f(x,y)dy$$

$$E(x.y) = E(x) + E(y)$$

Embora exista semelhança entre a média aritmética e ***E(x)***, a diferença é que ***E(x)*** é caracterizada como parâmetro associado a uma distribuição de probabilidade teórica.

4)Se (X,Y) é uma V.A. bidimencional e se X e Y forem independentes então:

$$E(x.y) = E(x)\,.\,E(y)$$

$$E(x.y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x\,.\,y\,.\,f(x,y)\,dx\,dy = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x\,.\,y\,.\,g(x).\,h(y).\,dx\,dy$$

$$E(x.y) = \int_{-\infty}^{+\infty} x\,.\,g(x)\,dx\,.\int_{-\infty}^{+\infty} y\,.\,h(y)\,dy = E(x).\,E(y)$$

Mediana *(md)*

É um ponto definido segundo a idéia de dividir a distribuição de probabilidade em duas partes equiprováveis.

$$P(X \langle\ md) = P(X \rangle\ md) = 0{,}5$$

Isso ocorre sempre no caso contínuo, em que a ***md*** pode também ser definida como ponto tal que **F(*md*) = 0,5**.

No caso discreto, quando a condição acima subsistir, haverá todo um intervalo cujos pontos satisfazem ela, convencionando-se em geral adotar o ponto médio desse intervalo.

Moda (m_o)

É ou são os pontos de maior probabilidade, no caso discreto, ou maior densidade de probabilidade, no caso contínuo. É portanto um parâmetro que indica a região mais provável da distribuição.

Variância (V(x) ou σ_x^2)

Definição

Seja **X** uma **VA**, definimos a variância de X, por:

$$Var(x) = \sigma^2 x = E[x - E(x)]^2$$

$$Var(x) = E[x^2] - [E(x)]^2$$

A raiz quadrada de V(x) é σx, denominado de desvio padrão

Teorema

Se $V(x) = E[x- E(x)]^2$, então $V(x) = E(x^2) - [E(x)]^2$

Exemplo

Em uma certa especulação comercial uma pessoa pode ter um lucro de 50.000,00 com probabilidade de 60% ou um prejuízo de 10.000,00 com probabilidade de 40%. Determine a expectativa de ganho dessa pessoa.

$$E(x) = \sum xi.\ Pi = 50.000 \times 0,4 + (-10.000) \times 0,6 = 26.000$$

$$V(x) = E(x^2) - [E(x)]^2$$

$$E(x^2) = (50.000)^2 \times 0,4 + (-10.000)^2 \times 0,6 = 2,5 \times 10^9 \times 0,4 + 1,0 \times 10^8 \times 0,6$$

$$E(x^2) = 1 \times 10^9 + 6 \times 10^7 = 1,06 \times 10^9$$

$$V(x) = 1,06 \times 10^9 - (26.000)^2 = 1,06 \times 10^9 - 6,76 \times 10^8$$

$$V(x) = 3,84 \times 10^8$$

Propriedades da Variância

1)Se "c" for uma constante

$$V(x + c) = V(x)$$

2) Se "c" for uma constante, então:

$$V(c \cdot x) = c^2 \cdot V(x)$$

3)Se (X,Y) for uma V.A. bidimencional e X e Y forem independentes, temos:

$$V(x + y) = V(x) + V(y)$$

4) Sejam, X_1, X_2, …, X_n V.A. independentes, então:

$$V(X_1 + X_2 + \ldots + X_n) = V(X_1) + V(X_2) + \ldots + V(X_n)$$

Teorema

Se X for uma VAC uniformemente distribuída sobre $[a, b]$, então:

$$V(x) = \frac{(b-a)^2}{12}$$

Soma de Variáveis Aleatórias

$\mu_{x+y} = \mu_x + \mu_y$ (a média da soma é a soma das médias)

$\sigma_{x \pm y} = \sqrt{\sigma^2_x + \sigma^2_y}$ (os desvios padrões nunca se somam ou se subtraem, somente as variâncias são aditivas

Observações:

- **X** é uma função e a denominamos **VA**;

- Nem toda função pode ser uma **VA**;

- Em alguns espaços amostrais o resultado **S** já constitui a característica numérica que desejamos registrar (***s* = *X* (*s*)**, neste caso teremos uma função identidade);

- Não nos interessa identificar a natureza funcional da função **X**, definida a **VA** interessa sim os valores possíveis de **X**.

Outras propriedades da esperança e da variância

1) Somando-se ou subtraindo-se uma constante (**K**) a cada valor de uma **VA**, a média fica somada ou subtraída desta constante, mas a variância não se altera;

$$\boxed{E(X \pm K) = E(X) \pm K}$$

Analogamente a variância de **X** será:

$$\boxed{Var(X \pm K) = Var(X)}$$

Partindo-se da variância de X, basta-se extrair a raiz quadrada para obtermos o desvio padrão.

2) Multiplicando-se ou dividindo-se cada valor de uma **VA** por uma constante, a média ficará multiplicada ou dividida pela constante; a variância ficará multiplicada ou dividida pelo quadrado da constante e o desvio padrão ficará multiplicado ou dividido pela constante:

$E(K.X) = K.E(X)$ $E(X/K) = E(X)/K$ com K ≠ 0	$Var(K.X) = K^2.Var(X)$ $Var(X/K) = Var(X)/K^2$ com K ≠ 0

3) A média da soma de duas variáveis aleatórias é igual a soma das médias destas variáveis. A média da diferença de duas variáveis é igual a diferença de suas médias

$$\boxed{E(X \pm Y) = E(X) \pm E(Y)}$$

4) A média de produto de variáveis aleatórias independentes é igual ao produto das mesmas variáveis.

$$\boxed{E(X.Y) = E(X).E(Y)}$$

4) A variância da soma de duas variáveis aleatórias independentes é igual a soma de suas variâncias. A variância da diferença de duas variáveis aleatórias independentes é igual a soma de suas variâncias.

$$\boxed{Var(X \pm Y) = Var(X) + Var(Y)}$$

6) Se cada uma das médias de duas variáveis aleatórias independentes é igual a zero, a média do quadrado da soma dessas variáveis é igual a soma das médias dos quadrados das mesmas variáveis.

$$\boxed{E[(X+Y)^2] = E(X^2) + E(Y^2)}$$