CAPÍTULO 4 - Variáveis aleatórias e distribuições de probabilidade

Conceito de variável aleatória

Uma função cujo valor é um número real determinado por cada elemento em um espaço amostral é chamado uma variável aleatória (v.a.)

Isso equivale a descrever os resultados de um experimento aleatório por meio de números ao invés de palavras, o que é uma grande vantagem pois possibilita melhor tratamento matemático, inclusive através de parâmetros que veremos a seguir.

obs.: Utilizaremos letras maiúsculas (X, Y, Z, etc.) para representar a v.a. e a correspondente letra minúscula (x, y, z, etc.) para um dos seus valores.

obs.: cada possível valor x de X representa um evento que é um subconjunto do espaço amostral.

exemplo:

Duas peças são retiradas sucessivamente, sem reposição, de uma caixa que contém quatro peças boas (B) e 3 defeituosas (D).

Para esse experimento teríamos o seguinte espaço amostral

S = {BB, BD, DB, DD}.

Se considerarmos a v.a. Y = número de peças boas retiradas teríamos:

S* = {0, 1, 2}.

Fazendo uma correspondência entre S e S* teríamos:

Evento Simples	Y = y
BB	2
BD	1
DB	1
DD	0

outro exemplo:

Considere o lançamento de duas moedas e seja X = nº de caras obtidas, c= cara e k=coroa

S = {cc, ck, kc, kk}
X = {0, 1, 2}.

Se um espaço amostral contém um numero finito de pontos como no exemplo anterior, ou uma sequência infinita enumerável de pontos amostrais, ele é chamado **espaço amostral discreto**. A v.a. definida sobre esse espaço é **chamada variável aleatória discreta** (v.a.d.).

Por outro lado, se um espaço amostral contém pontos amostrais que formam uma continuidade como todas as possíveis alturas de pessoas, pesos de animais, tempos de vida de um componente mecânico ou eletrônico, temperaturas, etc. então ele é chamado **espaço amostral contínuo**. A variável definida sobre esse espaço é chamada **variável aleatótia contínua** (v.a.c.).

obs.: na maior parte dos problemas práticos as v.a.c. representam dados medidos e as v.a.d. representam dados contados, tais como o número de defeituosos em uma amostra de n peças ou o número de acidentes na estrada Viçosa – Belo Horizonte no ano de 1999.

exemplo:

- nº de acidentes ocorridos em uma semana;
- nº de defeitos por peça produzida por um fabricante;
- nº de vitórias obtidas por um atleta;
- nº de filhos do sexo masculino por casal.

obs.: No caso das v.a.c. somente terão interesse as probabilidades de que a v.a. assuma valores em dados intervalos. As probabilidades de interesse poderão ser determinadas com o conhecimento da distribuição de probabilidade da v.a.

Distribuição de probabilidade

De acordo com o tipo de variável aleatória, ou seja se v.a.d. ou v.a.c., teremos:

a) X é uma v.a.d.

Nesse caso precisamos saber os valores da v.a.d. X e de sua função de probabilidade.

Chama-se **função de probabilidade** (f.p.) da variável aleatória discreta X, a função

$$P(X = x_i) = P(x_i) = p_i$$

que a cada valor de X (ou seja, a cada x_i) associa sua probabilidade de ocorrência.

obs.: essa função muitas vezes está reduzida a um valor numérico e, em outros casos, pode ser representada por uma fórmula.

A função $P(x_i)$ será uma função de probabilidade se satisfizer às seguintes condições:

i) $P(x_i) \geq 0$, para todo x_i

ii) $\sum_i P(x_i) = 1$

À coleção de pares $[x_i, P(x_i)]$, i = 1, 2, ..., n, denominaremos distribuição de probabilidade da v.a.d. X, que pode ser representada por meio de tabelas e gráficos.

exemplo:

Seja E: lançamento de um dado viciado de tal forma que a probabilidade é proporcional ao valor obtido no lançamento.

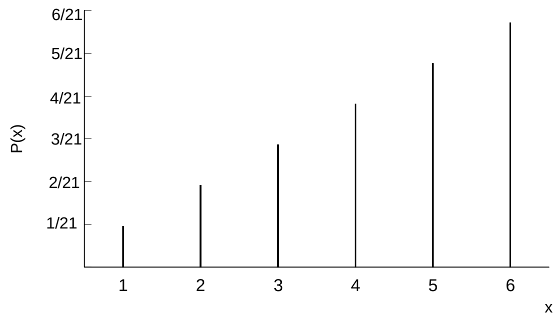
considere: S = {1, 2, 3, 4, 5, 6}

Vamos supor que estamos interessado na avaliação da variável aleatória X, onde X = {nº de pontos obtidos num lançamento}, ou X = {resultado num lançamento}

Assim os possíveis valores que a variável aleatória X pode assumir são:
X = {1, 2, 3, 4, 5, 6}, com as respectivas probabilidades:
$p + 2p + 3p + 4p + 5p + 6p = 1 \Rightarrow 21p = 1 \Rightarrow p = 1/21$

Para o nosso exemplo temos que a distribuição de probabilidades da v.a. X será:

i) Gráfico



ii) Tabela

x_i	1	2	3	4	5	6
$P(x_i)$	1/21	2/21	3/21	4/21	5/21	6/21

Observe que as duas condições apresentadas acima são satifeitas.

As tabelas e gráficos são utilizados para mostrar como a probabilidade total atribuída a um espaço amostral é distribuída pelos diversos resultados daquele espaço.

Nesse exemplo, a função de probabilidade poderia também ser representada como:

$$P(x_i) = \begin{cases} \dfrac{x_i}{21}, & \text{para } x_i = 1,2,3,4,5,6 \\ 0, & \text{para outros valores de } x_i \end{cases}$$

Problema proposto:

Uma urna contém 4 bolas azuis e 6 brancas. Duas bolas são retiradas sucessivamente I) com reposição e II) sem reposição. Determinar, em cada caso, a distribuição de probabilidade e a função de probabilidade do nº de bolas brancas retiradas.

b) X é uma v.a.c.

No caso de v.a.c. somente terão interesse as probabilidades de que a v.a. assuma valores em dados intervalos. Tais probabilidades poderão ser determinadas com o conhecimento da função densidade de probabilidade da v.a..

Chama-se **função densidade de probabilidade** (f.d.p.) da variável aleatória discreta X, a função f(x) que atenda às seguintes condições:

$$\text{i) } f(x) \geq 0, \qquad \text{para } a < x < b$$

$$\text{ii) } \int_a^b f(x)dx = 1,$$

onde a e b podem ser, respectivamente, $-\infty$ e ∞.

A f.d.p. assim definida determina a distribuição de probabilidades da v.a.c. X. Sua representação é dada em forma gráfica.

obs.:

1) Para $c < d$, $P(c < X < d) = \int_c^d f(x)\,dx$

2) Para um valor fixo de X, por exemplo, $X = x_0$, temos que $P(X = x_0) = \int_{x_0}^{x_0} f(x)dx = 0$;

sendo assim, as probabilidades abaixo são todas iguais, se X for uma v.a.c.:

$$P(c \leq X \leq d) = P(c \leq X < d) = P(c < X \leq d) = P(c < X < d)$$

3) A função densidade de probabilidade f(x), não representa probabilidade. Somente quando a função for integrada entre dois limites, ela produzirá uma probabilidade, que será a área sob a curva da função entre os valores considerados.

exemplo:
Uma v.a.c. X possui a seguinte função:

$$f(x) = \begin{cases} k, & \text{para } 0 \leq x < 1 \\ k(2-x), & \text{para } 1 \leq x < 2 \\ 0, & \text{para outros valores de } x \end{cases}$$

Pede-se:
a) A constante k para que f(x) seja uma f.d.p.
b) Gráfico da f(x).

c)$P(\frac{1}{2} \leq X \leq 1)$; d)$P(X = 1)$; e)$P(\frac{1}{2} \leq X < \frac{3}{2})$; f)$P(X > 2)$

Respostas: a) $2/3$; c) $1/3$; d) 0; e) $7/12$; f) 0.

Problemas propostos:

1. Seja uma v.a.c. X definida pela seguinte f.d.p.:

$$f(x) = \begin{cases} 0, & \text{para } x < 0 \\ kx, & \text{para } 0 \leq x \leq 2 \\ 0, & \text{para } x > 2 \end{cases}$$

a) Determinar o valor de k
b) Traçar o gráfico da f.d.p.
c) Calcular $P(X \leq 1)$.

Respostas: a) $1/2$; c) $1/4$

2. Uma v.a.c. X tem a seguinte f.d.p.

$$f(x) = \begin{cases} 0, & \text{para } x < 0 \\ kx, & \text{para } 0 \le x < 5 \\ k(10 - x), & \text{para } 5 \le x \le 10 \\ 0, & \text{para } x > 10 \end{cases}$$

a) Determinar o valor de k
b) Traçar o gráfico da f.d.p.
c) Calcular $P(X \geq 3)$

Respostas: a) $1/25$; b) $41/50$

3. Considere a seguinte função de X

$$f(x) = \begin{cases} k\,e^{-3x} & \text{para } x > 0 \\ 0 & \text{c. c.} \end{cases}$$

a) Encontre o valor de k para que f(x) seja uma função densidade de probabilidade;
b) Encontre $P(0{,}5 \le X \le 1)$;
c) Faça o gráfico da f(x).

Respostas: a) k = 3; b) 0,173

Existem muitos problemas nos quais é de interesse conhecer a probabilidade que a v.a. X assumisse valores menores que um particular valor x.

Nesse caso precisamos definir a função de distribuição acumulada de X.

Função de distribuição acumulada

Dada a variável aleatória X, chamaremos de função de distribuição acumulada ou, simplesmente, função de distribuição F(x) a função $F(x) = P(X \le x)$.

Observe que o domínio de F é todo o conjunto real.

obs.:

a) $0 \le F(x) \le 1$ para todo x.
b) Se $x_1 \le x_2$, então $F(x_1) \le F(x_2)$, isto é, F(x) é não-decrescente.

a) F(x) para X v.a.d.

Para X uma v.a.d. temos que:

$$F(x) = P(X \le x) = \sum_{t \le x} P(t)$$

Exemplo: Seja X uma v.a.d. com a seguinte distribuição de probabilidade

x_i	0	1	2	3	4	total
$P(x_i)$	1/16	4/16	6/16	4/16	1/16	1,00

Pede-se:

a) Traçar o gráfico da distribuição de probabilidade de X.
b) Obter a função de distribuição acumulada e traçar seu gráfico.

b) F(x) para X v.a.c.

Para X uma v.a.c. temos que:

$$F(x) = P(X \le x) = P(-\infty < X \le x) = \int_{-\infty}^{x} f(t)dt$$

Temos ainda que,

$$P(c < X \le d) = F(d) - F(c) = \int_{c}^{d} f(x)dx\,.$$

Vê-se também que $f(x) = \dfrac{d\,F(x)}{dx}$ em todos os pontos de continuidade de f(x), isto é, a derivada da função de distribuição é a função densidade de probabilidade.

Exemplos:
1. Seja X uma v.a.c. com a seguinte f.d.p.:

$$f(x) = \begin{cases} \dfrac{1}{2}x, & \text{para } 0 \le x \le 2 \\ 0, & \text{caso contrario.} \end{cases}$$

Pede-se:
a) Traçar o gráfico da f.d.p.
b) Obter F(x) e traçar seu gráfico
c) Calcular $P(X \le 1)$
d) Calcular $P\left(1 \le X \le \dfrac{3}{2}\right)$
e) Qual o valor de X abaixo do qual existem 50% dos dados?

Respostas: c) $1/4$; d) $5/16$; e) $\sqrt{2}$

2. Seja X uma v.a.c. com função dada por:

$$f(x) = \begin{cases} ax, & \text{se } 0 \le x < 1 \\ a, & \text{se } 1 \le x < 2 \\ -ax + 3a, & \text{se } 2 \le x \le 3 \\ 0, & \text{se } x < 0 \text{ ou } x > 3 \end{cases}$$

Pede-se:
a) Determine a constante <u>a</u> para que f(x) seja uma f.d.p.
b) Traçar o gráfico da f.d.p.
c) Obter F(x) e traçar seu gráfico
d) $P\left(0 < X < \dfrac{3}{2}\right)$

e) Se X_1, X_2 e X_3 forem três observações independentes de X, qual será a probabilidade de, exatamente um desses três números ser maior do que 1?
f) Qual valor de X abaixo do qual existem 50% dos dados?

Resposta: a) $1/2$; d) $1/2$; e) $9/64$; f) 1,5.

Distribuições empíricas

Geralmente quando temos um experimento envolvendo uma v.a.c. a verdadeira função densidade (f(x)) nos é desconhecida. Para que adotemos uma f(x) razoável uma boa análise sobre todas as informações disponíveis se faz necessária. Essa análise pode ser feita com o uso de distribuições de frequências relativas, em forma de tabelas ou gráficos, conforme descrito no capitulo 2. Distribuições assim obtidas permitem supor acerca da “verdadeira” distribuição de probabilidades associada àquela variável aleatória. Supondo-se então uma determinada distribuição de probabilidades à v.a. em questão, outras análises mais avançadas podem ser realizadas (como os testes de hipóteses para se testar tais suposições). Exemplos desse tipo serão citados oportunamente.

Distribuições de probabilidades conjuntas

Vimos, anteriormente o caso de distribuições de probabilidades para o caso de uma única variável aleatória. Há casos, contudo, nos quais estamos interessados no estudo de duas ou mais variáveis aleatórias simultaneamente. Por exemplo, alguém poderia estar interessado na dureza (D) e na tensão de ruptura (T) de uma liga de cobre fornecendo assim o resultdo (b,t).

Se X e Y são duas v.a., a distribuição de probabilidade para a sua ocorrência simultânea pode ser representada por f(x,y) para o caso em que (X,Y) é uma v.a. bidimensional contínua, e p(x,y) para o caso em que (X,Y) é uma v.a. bidimensional discreta.

Definição

Sejam E um experimento e S um espaço amostral associado a E.

Sejam X = X(s) e Y =Y(s), duas funções, cada uma associando um número real a cada resultado $s \in S$. Denominaremos (X,Y) uma variável aleatória bidimensional.

Em determinadas situações, X e Y não estão necessariamente ligadas a um mesmo experimento, mas existe uma razão bastante definida para considerar X e Y conjuntamente.

Para o nosso estudo vamos considerar que X e Y são ambas discretas ou contínuas.

Do mesmo modo que no caso unidimensional (X,Y) deve ter associada, a cada valor que pode assumir, uma probabilidade de sua ocorrência. Assim precisamos definir a distribuição de probabilidade da v.a. bidimensional (X,Y).

Distribuição conjunta de duas variáveis aleatórias, distribuições marginais e condicionais

a) <u>(X, Y) é v.a.d. bidimensional</u>

(X,Y) será uma v.a.d. bidimensional se os valores possíveis de X e Y forem finitos ou infinitos enumeráveis. Isto é, se os valores possíveis de (X,Y) podem ser representados por $(x_i, y_j) \quad i = 1, 2, \ldots, r \quad \text{e} \quad j = 1, 2, \ldots, s$.

a.1) Função de probabilidade conjunta de X e Y

Chama-se de função de probabilidade conjunta da v.a.d.bidimensional (X, Y) a função:

$$P(X{=}x_i\,;\, Y{=}y_j) = P(x_i;\, y_j) = p_{ij}$$

que a cada valor de $(x_i; y_j)$ associa sua probabilidade de ocorrência.

Para que $P(x_i; y_j)$ seja uma função de probabilidade conjunta é necessário que satisfaça às seguintes condições:

$$\text{a) } P(x_i, y_j) \geq 0, \quad \text{para todo par } (x_i, y_j)$$

$$\text{b) } \sum_i \sum_j P(x_i, y_j) = 1$$

Distribuição de probabilidade conjunta

É o conjunto $\{(x_i; y_j), P(x_i; y_j)\}$ para i = 1, 2,..., r e j = 1, 2,..., s

a.2) Distribuições marginais

Dada uma distribuição conjunta de duas variáveis aleatórias X e Y, podemos determinar a distribuição de X sem considerar Y e a de Y sem considerar X. São as chamadas distribuições marginais.

A distribuição marginal é constituida pelos valores da variável aleatória e suas respectivas probabilidades marginais, que serão apresentadas na forma gráfica ou tabular conforme visto para v.a. unidimensionais. A probabilidade marginal para cada valor é obtida da seguinte forma:

$$-\text{ para X:} \quad P(X = x_i) = P(x_i) = \sum_{j=1}^{s} P(x_i, y_j)$$

$$-\text{ para Y:} \quad P(Y = y_j) = P(y_j) = \sum_{i=1}^{r} P(x_i, y_j)$$

a.3) Distribuições condicionais

Seja x_i um valor de X, tal que $P(X = x_i) = P(x_i) > 0$

A probabilidade

$$P(Y = y_j / X = x_i) = \frac{P(X = x_i, Y = y_j)}{P(X = x_i)}$$

é denominada probabilidade condicional de $Y = y_j$, dado que $X=x_i$.

Assim, para x_i fixado, os pares $[y_j, P(Y = y_j / X = x_i)]$ definem a distribuição condicional de Y, dado que $X = x_i$.

Analogamente para o X:

$$P(X = x_i / Y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)}$$

b) <u>(X, Y) é v.a.c. bidimensional</u>

A variável (X, Y) será uma v.a.c.bidimensional se (X, Y) puder assumir todos os valores em algum conjunto não enumerável.

b.1) Função densidade de probabilidade conjunta

Seja (X, Y) uma v.a.c.bidimensional. Dizemos que f(x, y) é uma função densidade de probabilidade conjunta de X e Y, se satisfizer às seguintes condições:

a) $f(x, y) \geq 0$, para todo (x, y)

b) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)dxdy = 1$

$f(x,y) = 0$ para $(x,y) \notin$ aos intervalos de x e y.

Temos ainda que:

$$P(a \leq X \leq b, c \leq Y \leq d) = \int_c^d \int_a^b f(x, y)dxdy$$

b.2) distribuições marginais

As f.d.p. marginais de X e Y são dadas por:

$$g(x) = \int_{-\infty}^{\infty} f(x, y)dy \quad \text{e} \quad h(y) = \int_{-\infty}^{\infty} f(x, y)dx \quad \text{respectivamente.}$$

Temos ainda que:

$$P(a \leq X \leq b) = \int_a^b g(x)dx \quad \text{e} \quad P(c \leq Y \leq d) = \int_c^d h(y)dy$$

b.3) Distribuições condicionais

Sejam X e Y v.a.c. com f.d.p. conjunta f(x, y) e f.d.p. marginais dadas por g(x) e h(y).

A f.d.p. condicional de X, dado que Y = y é definida por:

$$f(x / y) = \frac{f(x, y)}{h(y)}, \quad h(y) > 0$$

Analogamente, a f.d.p. condicional de Y, dado que X = x é definida por:

$$f(y/x) = \frac{f(x,y)}{g(x)}, \quad g(x) > 0$$

As f.d.p. condicionais acima, satisfazem a todas as condições impostas para uma f.d.p. unidimensional.

Deste modo para y fixado, teremos:

$$a) f(x/y) \geq 0$$

$$b) \int_{-\infty}^{\infty} f(x/y)dx = \int_{-\infty}^{\infty} \frac{f(x,y)}{h(y)}dx = \frac{1}{h(y)} \int_{-\infty}^{\infty} f(x,y)dx = \frac{h(y)}{h(y)} = 1$$

Variáveis aleatórias independentes

a) <u>(X, Y) é uma v.a.d.bidimensional.</u>

Definição 1 - Seja (X, Y) v.a.d.bidimensional. Dizemos que X e Y são independentes se, e somente se, para todo par de valores (x_i, y_j) de X e Y, tem-se:

$$P(X = x_i, Y = y_j) = P(X = x_i).P(Y = y_j)$$

Basta que esta condição não se verifique para um par (x_i, y_j) para que X e Y não sejam independentes. Neste caso diremos que X e Y são dependentes.

Definição 2 - Seja (X, Y) uma v.a.d.bidimensional. Neste caso X e Y serão independentes se, e somente se,

$$P(X = x_i / Y = y_j) = P(X = x_i), \text{ para todo i e j.}$$

ou equivalente se, e somente se:

$$P(Y = y_j / X = x_i) = P(Y = y_j), \text{ para todo i e j.}$$

b) <u>(X, Y) é uma v.a.c.bidimensional.</u>

Definição 1 - Seja (X, Y) v.a.c.bidimensional. Diremos que X e Y são independentes se, e somente se:

$$f(x, y) = g(x).h(y), \text{ para todo x e todo y}$$

Definição 2 - Seja (X, Y) uma v.a.c.bidimensional. Neste caso X e Y serão independentes se, e somente se:

$f(x/y) = g(x)$. Nesse caso, é evidente que $f(y/x) = h(y)$.

OBS.: Todas as definições concernentes a duas v.a. podem ser generalizadas para o caso de n v.a.

Exemplos:

1. Dada a distribuição de probabilidade conjunta de (X, Y) na Tabela abaixo

X \ Y	0	1	2
0	0,10	0,20	0,20
1	0,04	0,08	0,08
2	0,06	0,12	0,12

Pede-se:
a) Distribuição marginal de X e Y, distribuição de (X + Y) e de XY;
b) X eY são independentes ?
c) As distribuições condicionais de X dado que Y = 0 e Y dado que X = 1.
d) $P(X \geq 1, Y \leq 1)$
e) $P(X \leq 1 / Y = 0)$
f) Construir a distribuição conjunta a partir das distribuições marginais de X e de Y.

Resposta: d) 0,30; e) 0,70

2. Sejam X e Y v.a.c. com f.d.p. conjunta dada por:

$$f(x,y) = \begin{cases} k(2x+y), & 2 \leq x \leq 6, \quad 0 \leq y \leq 5 \\ 0, \text{ para outros valores de } x \text{ e } y \end{cases}$$

Pede-se:
a) O valor de k
b) $P(X \leq 3,\ 2 \leq Y \leq 4)$
c)P(Y < 2)
d) P(X > 4)
e) X e Y são v.a. independentes? Justifique
f) f(x/ y)
g) f(x/ Y = 1)

Resposta: a) $1/210$; b) $16/210$; c) $72/210$; d) $125/210$; e) Não

Problemas propostos:
1) A função de probabilidade conjunta da v.a.X e Y é dada por

$$p(x,y) = \frac{10!}{x!\,y!\,(10-x-y)!}\left(\frac{1}{36}\right)^x\left(\frac{25}{36}\right)^y\left(\frac{10}{36}\right)^{10-x-y}$$

para x = 0,1,2, …, 10
y = 0,1,2, …, 10
$0 \leq x + y \leq 10$

Calcule: a) P(X ≤ 2, Y ≤ 2); b) P(X + Y = 4)

Resposta: a) ≅ 0,00157; b) ≅ 0,0262

2) Sejam X, Y e Z três v.a. independentes e cada uma com a função densidade

$$f(t) = \begin{cases} e^{-t} & \text{para } t \geq 0 \\ 0 & \text{c.c.} \end{cases}$$

Ache a P(X ≤ 3, 0,5 ≤ Y ≤ 2,5, Z > 1)
Resposta: ≅ 0,18332

Conceitos e Propriedades de esperança matemática e variância de variáveis aleatórias

a) Esperança matemática (média ou valor esperado de uma v. a.)

Neste item vamos aprender a quantificar o parâmetro média de uma população. A esperança matemática de uma população é também denominada uma medida da tendência central.

Parâmetro é uma medida utilizada para descrever uma característica de uma população e caracteriza a distribuição de probabilidade de uma variável aleatória.

Sob o ponto de vista científico, a esperança matemática corresponde ao que se espera que aconteça em média.

a.1) Caso em que X é uma v.a.d.

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

x_i	x_1	x_2	...	x_n	Total
$P(x_i)$	$P(x_1)$	$P(x_2)$	...	$P(x_n)$	1

Define-se a esperança matemática de X por:

$$E(X) = \mu_x = \mu = x_1 . P(x_1) + x_2 . P(x_2) + \ldots + x_n . P(x_n)$$

$$E(X) = \sum_{i=1}^{n} x_i . P(x_i)$$

Exemplo: Um fabricante produz peças tais que 10% delas são defeituosas e 90% não são defeituosas. Se uma peça defeituosa for produzida, o fabricante perde U\$ 1,00, enquanto uma peça não defeituosa lhe dá um lucro de U\$ 5,00. Seja a variável aleatória X = {lucro líquido por peça}. Calcular a média do lucro líquido por peça.

Resposta: $E(X) = 4{,}40$

a.2) Caso em que X é uma v.a.c.

A esperança matemática de uma v.a.c. X é definida por:

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

Exemplo: Uma v.a.c. X apresenta a seguinte f.d.p.:

$$f(x) = \begin{cases} 0, & \text{para } x < 0 \\ \dfrac{x}{2}, & \text{para } 0 \le x \le 2 \\ 0, & \text{para } x > 2 \end{cases}$$

Calcular E(X)

Resposta: $4/3$

a.3) Propriedades da esperança matemática

As propriedades a seguir apesar de estarem apenas demonstradas para quando X é uma v.a.c., valem também para quando X é uma v.a.d..

P_1) Se X é uma v.a. com P(X = k) = 1, então E(X) = k, sendo k uma constante (ou numa linguagem mais simples mas menos rigorosa, pode-se dizer que a média de uma constante é a própria constante).

<u>Prova</u>: $E(X) = \int_{-\infty}^{\infty} kf(x)dx = k\int_{-\infty}^{\infty} f(x)dx = k \quad \rightarrow \quad E(X) = k$

P_2) A esperança matemática do produto de uma constante por uma variável é igual ao produto da constante pela esperança matemática da variável, ou seja, multiplicando-se uma variável aleatória por uma constante, sua média fica multiplicada por essa constante.

<u>Prova</u>: $E(kX) = \int_{-\infty}^{\infty} kxf(x)dx = k\int_{-\infty}^{\infty} xf(x)dx = kE(X)$

P_3) A esperança matemática do produto de duas variáveis aleatórias <u>independentes</u> é igual ao produto das esperanças matemáticas das variáveis, ou seja, a média do produto de duas variáveis aleatórias <u>independentes</u> é o produto das médias.

<u>Prova</u>: $E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy$

Se X e Y são v.a. independentes, a f.d.p. conjunta pode ser fatorada no produto das f.d.p. marginais de X e Y. Assim:

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyg(x)h(y)dxdy = \int_{-\infty}^{\infty} xg(x)dx\int_{-\infty}^{\infty} yh(y)dy$$

$$E(XY) = E(X)E(Y)$$

obs: E(XY) = E(X).E(Y) não implica que X e Y sejam variáveis aleatórias independentes.

P_4) A esperança matemática da soma ou da subtração de duas v.a. quaisquer é igual a soma ou a subtração das esperanças matemáticas das duas v.a., ou seja, a média da soma ou da subtração de duas v.a. é igual a soma ou subtração das médias.

<u>Prova</u>:

$$E(X \pm Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x \pm y)\, f(x,y)dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, f(x,y)dxdy \pm \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\, f(x,y)dxdy$$

$$E(X \pm Y) = \int_{-\infty}^{\infty} x\, g(x)dx \pm \int_{-\infty}^{\infty} y\, h(y)dy = E(X) \pm E(Y)$$

P_5) A esperança matemática da soma ou subtração de uma v.a. com uma constante é igual a soma ou subtração da esperança matemática com a constante, ou seja, somando-se ou subtraindo-se uma constante a uma v.a., a sua média fica somada ou subtraída da mesma constante.

<u>Prova</u>: $E(X \pm K) = \int_{-\infty}^{\infty} (x \pm K)f(x)dx = \int_{-\infty}^{\infty} xf(x)dx + \int_{-\infty}^{\infty} Kf(x)dx = E(X) \pm K$

P_6) A média de uma v.a. centrada é zero, ou seja, a média dos desvios dos valores da v.a. em relação a sua média é zero.

Obs: Dizemos que a v.a. está centrada quando todos os valores são expressos como desvios em relação à respectiva média, $(X - \mu_x)$.

Assim: $E[X - \mu_x] = E(X) - E[\mu_x] = \mu_x - \mu_x = 0$

b) Variância

É a medida que quantifica a dispersão dos valores em torno da média.

A variância de uma v.a. X é definida por:

$$V(X) = \sigma_x^2 = E[X - E(X)]^2 = E[X - \mu_x]^2$$

b.1.) Para X v.a.d.: $V(X) = \sum_i (x_i - \mu_x)^2 P(x_i)$

b.2.) Para X v.a.c.: $V(X) = \int_{-\infty}^{\infty} (x - \mu_x)^2 f(x)dx$

Uma fórmula mais prática para calcular a variância é:

$$V(X) = E(X^2) - [E(X)]^2,$$

pois,

$$\begin{aligned} V(X) &= E[X - E(X)]^2 \\ &= \{X^2 - 2XE(X) + [E(X)]^2\} \\ &= E(X^2) - 2E(X)E(X) + [E(X)]^2 \\ &= E(X^2) - [E(X)]^2 \end{aligned}$$

em que:

- para X v.a.d.: $E(X^2) = \sum_i x_i^2 P(x_i)$

- para X v.a.c.: $E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)dx$

b.3.) Propriedades da variância:

Valem tanto para X v.a.d. quanto para X v.a.c.

P_1) A variância de uma constante é igual a zero.

Prova: $V(k) = E[K - E(K)]^2 = E(k - k)^2 = 0$

P_2) Somando-se ou subtraindo-se uma constante à uma v.a., sua variância não se altera.

Prova:

$$V(X \pm k) = E[(X \pm k) - E(X \pm k)]^2$$
$$= E[(X) - E(X) \pm (k - k)]^2$$
$$= E[X - E(X)]^2 = V(X)$$
$$V(X \pm k) = V(X)$$

P_3) Multiplicando-se uma v.a. por uma constante, sua variância fica multiplicada pelo quadrado da constante.

Prova:

$$V(kX) = E[kX - E(kX)]^2$$
$$= E[kX - kE(X)]^2$$
$$= k^2 E[X - E(X)]^2$$
$$V(kX) = k^2 V(X)$$

P_4) A variância da soma de duas v.a. independentes é igual a soma das variâncias das duas variáveis.

Prova:

$$V(X+Y) = E[X+Y]^2 - [E(X+Y)]^2$$
$$= E(X^2 + 2XY + Y^2) - [E(X) + E(Y)]^2$$
$$= E(X^2) + 2E(XY) + E(Y^2) - \{[E(X)]^2 + 2E(X)E(Y) + [E(Y)]^2\}$$
$$= E(X^2) + 2E(XY) + E(Y^2) - [E(X)]^2 - 2E(X)E(Y) - [E(Y)]^2$$

Se X e Y são independentes; E(XY) = E(X).E(Y), assim

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

Do mesmo modo:

$$V(X - Y) = V(X) + V(Y)$$

OBS.1: Desvio padrão da variável X é a raiz quadrada positiva da variância de X.

$$\sigma_x = \sqrt{V(X)}$$

OBS.2: Sejam X e Y duas v.a.. A covariância entre X e Y, denotada por Cov(X,Y) é definida por:

Cov(X,Y) = E{[X – E(X)][Y – E(Y)]}, ou, de forma alternativa,

Cov(X,Y) = E(X,Y) – E(X)E(Y)

A covariância nos dá uma idéia da relação de dependência entre duas ou mais v.a.

Proposições úteis:

a) Cov(X,Y) = Cov(Y,X)
b) Se V(X) = 0 ou V(Y) = 0 então a Cov(X,Y) = 0
c) Cov(X,k) = Cov(k,Y) = 0
d) Cov(aX,bY) = abCov(X,Y), sendo a e b constantes
e) Cov(X + Z,Y) = Cov(X,Y) + Cov(Z,Y)

Resultado útil: Se X e Y são duas variáveis quaisquer

$$V(aX \pm bY) = a^2V(X) + b^2V(Y) \pm 2abCov(X,Y)$$

Exercícios propostos:

1. Sabendo-se que Y=3X-5 e que E(X)=2 e V(X)=1, calcule:

a)E(Y); b)V(Y); c)E(X+3Y); d)$E(X^2 + Y^2)$; e)V(3X+2Y);

Resp.: 1; 9; 5; 15; 81

2. Uma urna contém 5 bolas brancas e 7 bolas pretas. Três bolas são retiradas simultaneamente dessa urna. Se ganharmos R$ 200,00 por bola branca retirada e perdermos R$ 100,00 por bola preta retirada, qual seria o nosso lucro esperado? Resp.: 75

3. Uma moeda honesta é lançada sucessivamente até sair cara ou até serem feitos 3 lançamentos. Obtenha a distribuição de X = número de lançamentos, e calcule sua média e variância. Resp.: E(X)=1,75; Moda=1; V(X)=11/16

4. Uma máquina de apostar tem 2 discos independentes. Cada disco tem 10 figuras: 4 maçãs, 3 bananas, 2 pêras e 1 laranja. Paga-se 80 para acionar a máquina. Se aparecerem 2 maçãs ganha-se 40; 2 bananas 80; 2 pêras 140 e 2 laranjas 180. Qual é o resultado esperado após inúmeras jogadas? Resp.: E(X) = – 59

5. Um determinado artigo é vendido em caixas a preço de 8 U.M. por caixa. Sabe-se que 20% dos artigos vendidos apresentam algum defeito de fabricação. Um comprador faz a seguinte proposta: Pede para poder amostrar, ao acaso, 10 artigos por caixa e pagará por caixa 10 U.M. se nenhum dos artigos amostrados for defeituoso; 5 U.M. se um ou dois artigos forem defeituosos e 4 U.M. se três ou mais forem defeituosos.

O que é melhor para o vendedor; manter o seu preço de 8 U.M. por caixa ou aceitar a proposta do comprador? Mostre por quê.

Considere X = número de artigos defeituosos, com a seguinte distribuição de probabilidade: (**sugestão**: utilize a variável Y = valor pago por caixa)

x_i	0	1	2	≥ 3	Total
$P(x_i)$	0,1074	0,2684	0,3019	0,3223	1,00

Resp.: O vendedor deve manter o seu preço [$E(Y) \cong 5,21$]

6. (X, Y) é uma variável aleatória bidimensional discreta com a seguinte distribuição conjunta:

X \ Y	-3	2	4	$P(x_i)$
1	0,1	0,2	0,2	0,5
3	0,3	0,1	0,1	0,5
$P(y_j)$	0,4	0,3	0,3	1

Pede-se calcular:

a) E(X), V(X) e σ_x

b) E(Y), V(Y) e σ_y

c) E(X + Y)

d) X e Y são independentes ?

Resp.: a) 2; 1 e 1 b) 0,6; 9,24 e 3,04 c) 2,6 d) não são

7. Seja X uma v.a.c. com a seguinte f.d.p.:

$$f(x) = \begin{cases} \frac{2}{3}, & x \in [0,1] \\ \frac{2}{3}(2-x), & x \in [1,2] \\ 0, & \text{caso contrá ric} \end{cases}$$

Calcule:
a) A esperança matemática de $(X-1)^2$
b) O desvio padrão de X
Resp.: a) 0,278 b) 0,478

8. Mostre que E{[X – E(X)][Y – E(Y)]}= E(X,Y) – E(X)E(Y)

Universidade Federal de Viçosa - CCE / DPI
INF 161 - Iniciação à Estatística / Inf 162 – Estatística I
Lista de Exercícios: Cap. 4 - Variáveis Aleatórias

1. Uma v.a.c. X possui f.d.p. dada por :

$$f(x) = \begin{cases} 6(x - x^2) \text{ , } 0 \le x \le 1 \\ \quad 0 \qquad \text{, para outros valores de x} \end{cases}$$

Calcular :
a) $P\left[E(X) - 2\sigma \le X \le E(X) + 2\sigma\right], \sigma = \sqrt{V(X)}$, dado $E(X^2) = 3/10$
b) F(x), a Função de distribuição acumulada

2. Suponha que X e Y tenham a seguinte distribuição conjunta :

Y \ X	-1	1	2	
-2	0,1	0,1	0	
0	0,1	0,2	0,3	
3	0,1	0,1	0	

Sabendo-se que V(X) = 1,41 e E(Y) = 0,2 ; pede-se:
a) X e Y são v. a. independentes? Mostre.
b) $E\left(\frac{1}{2}X - 3Y^2 + 8\right)$
c)$V\left(\frac{1}{2}X - 3Y^2 + 8\right)$

3. Dada a função densidade de probabilidade abaixo:

$$f(x) = \begin{cases} \frac{1}{2} & \text{, se } 0 \le x \le 1 \\ -\frac{1}{4}(x-3) & \text{, se } 1 \le x \le 3 \\ 0 & \text{, para outros valores de x} \end{cases}$$

Calcule :

a) V(12X - 8), dado $E(X^2) = 127/6$

b) A função de distribuição acumulada

c) P(0,5 < X < 1,5)

4. Uma v. a. c. X possui a seguinte f. d. p. :

$$f(x) = \begin{cases} 0 & \text{, se } x < -1 \text{ ou } x \ge 4/3 \\ \frac{1}{2}(1-x^2) & \text{, se } -1 \le x < 0 \\ \frac{1}{2} & \text{, se } 0 \le x < 4/3 \end{cases}$$

Determinar:

a) F(x), a função de distribuição acumulada

b) $P(-0{,}5 \le x \le 0{,}5)$

5.Dada a distribuição conjunta abaixo, parcialmente indicada:

Y \ X	-3	-2	-1	P(Y)
-2	1/15	1/15	?	7/30
0	8/30	?	2/15	?
1	?	1/30	?	7/30
P(X)	6/15	7/30	?	

Pede-se:

a) Verifique se X e Y são v. a. independentes.

b) $E\left(\frac{X^2}{3} - \frac{2Y}{5} - 10\right)$

c) $V(8 - 15X)$

6.Cite as propriedades de:

a) Esperança Matemática.

b) Variância.

7. Conceitue:

a) Variável aleatória discreta

b) Variável aleatória contínua

8.Uma v. a. c. é dada por:

$$f(x) = \begin{cases} kx^2 & , \text{se } 2 \le x < 5 \\ k(8-x) & , \text{se } 5 \le x \le 8 \\ 0 & , \text{se x assume outros valores} \end{cases}$$

Determinar:
a) O valor da constante k para que f(x) seja uma f. d. p.
b) $P\left(\frac{8}{2} \le X \le \frac{27}{2}\right)$

9. Suponha que X e Y tenham a seguinte distribuição conjunta:

Y \ X	1	2	4	5	Soma
2	0,2	0,1	0,1	0,2	
3	0,1	0,1	0,1	0,1	
Soma					

Pede-se:
a) $E\left(-\frac{1}{3}X\right)$; b) $V(5X-3Y)$; c) X e Y são v. a. independentes? Mostre porque.

10. Sabendo-se que X e Y são variáveis aleatórias independentes e que E(X) = 5, V(X) = 2, E(Y) = 8 e V(Y) = 3, calcule:

a) E(X - Y + 3) d) $V(3Y+2)$

b) E[(X - Y)^2] c) $V\left(X - \frac{1}{3}Y\right)$

11. Seja (X, Y) uma variável aleatória bidimensional discreta, com a seguinte função de probabilidade:

$$P(x_i, y_j) = \begin{cases} \frac{2x_i + y_j}{42} & , \text{para } x_i = 0, 1, 2 \text{ e } y_j = 0, 1, 2, 3 \\ 0 & , \text{caso contrá rio} \end{cases}$$

Pede-se:
a) Dar a tabela da distribuição de probabilidade conjunta.
b) Dar a tabela da distribuição marginal de X e também a de Y.
c) $E(X - 2Y + 4)$

12. Dada a seguinte função:

$$f(x) = \begin{cases} K(2-x) & , \text{se } 0 \le x < 1 \\ K & , \text{se } 1 \le x \le 2 \\ 0 & , \text{caso contrá rio} \end{cases}$$

Determinar:
a) O valor de K para que f(x) seja uma f.d.p.
b) $E(X^3)$
c) $P\left(X \geq \frac{3}{2}\right)$ d) $P(X = 1)$

14. Sejam X e Y v. a. c. com função densidade de probabilidade conjunta dada por:

$$f(x,y) = \begin{cases} K(2x+y), \text{ se } 2 \leq x \leq 6 \quad \text{e} \quad 0 \leq y \leq 5 \\ 0 \qquad , \text{c. c.} \end{cases}$$

Pede-se:
a) O valor de K para que $f(x,y)$ seja uma f. d. p.
b) $P(Y \leq 2)$
c) $P(X > 3 / 0 \leq Y \leq 2)$
d) X e Y são v. a. independentes? mostre.

15.A variável aleatória contínua X tem f. d. p., f(x) = $3x^2$, $-1 \leq x \leq 0$. Se a for um número que satisfaça a $-1 \leq a \leq 0$, calcule:
$P\left(X > a \;/\; X < \frac{a}{2}\right)$

16. Dado $f(x,y) = \begin{cases} \frac{3-y}{16}, \text{ se } 0 \leq y \leq 2 \quad \text{e} \quad 0 \leq x \leq 4 \\ 0 \qquad , \text{caso contrário} \end{cases}$

Determinar:
a) As funções marginais de X e Y
b) Se X e Y são v. a. independentes.

17. Seja f(x, y) = 2(x + y - 2xy), para $0 \leq x \leq 1$ e $0 \leq y \leq 1$ e f(x, y) = 0, para quaisquer outros valores de x e de y. Pede-se:
a) Mostre que f(x, y) é uma f.d.p.
b) Obtenha a f.d.p. marginal de X e a de Y.

18. Suponha que as dimensões, X e Y, de uma chapa retangular de metal, possam ser consideradas variáveis aleatórias contínuas independentes, com as seguintes f.d.p.:

$$g(x) = \begin{cases} x-1, \; 1 < x \leq 2 \\ -x+3, \; 2 < x < 3 \\ 0 \quad , \text{ para outros valores} \end{cases} \quad \text{e} \quad h(y) = \begin{cases} \frac{1}{2}, \; 2 < y < 4 \\ 0 \; , \text{ para outros valores} \end{cases}$$

Pede-se: Ache a f.d.p. da área da chapa (A).

RESPOSTAS

1. a) ≅ 0,9855 b) 0 se $x < 0$; $x^2(3-2x)$ se $0 \leq x < 1$; 1 se $x \geq 1$

2. a) não b) 0,55 c)119,57

3. a) 2879 b) $0 \text{ se } x < 0; \quad \frac{x}{2} \text{ se } 0 \leq x < 1; \quad -\frac{1}{8}(x^2 - 6x + 1) \text{ se } 1 \leq x < 3; \; 1 \text{ se } x \geq 3$
c) 15/32

4.a) $0 \text{ se } x < -1; \; \frac{1}{6}(-x^3 + 3x + 2) \text{ se } -1 \leq x < 0; \quad \frac{1}{6}(2 + 3x) \text{ se } 0 \leq x < 4/3; \quad 1 \text{ se } x \geq 4/3$
b) 23/48

5. a) não b) -3723/450 c) 689/4

8. a) 2/87 b) 149/261

9. a) -1 b) 72,16 c) não

10. a) 0 b) 14 c) 7/3 d) 27

11. c) 70/42

12. a) 2/5 b) 81/50 c) 0,2 d) 0

14.a) 1/210 b) 72/210 c) 60/72 d) não

15. $\frac{-7a^3}{\left(a^3 + 8\right)}$

16. a) $g(x) = \begin{cases} \frac{1}{4}, & 0 \leq x \leq 4 \\ 0, & \text{c.c.} \end{cases} \qquad h(y) = \begin{cases} \frac{1}{4}\left(3 - y\right), & 0 \leq y \leq 2 \\ 0, & \text{c.c} \end{cases}$ b) sim

17. a) $\int_0^1 \int_0^1 f(x,y)\,dxdy = 1$ b) $g(x) = \begin{cases} 1, & 0 \leq x \leq 1 \\ 0, & \text{c.c.} \end{cases}$ e $h(y) = \begin{cases} 1, & 0 \leq y \leq 1 \\ 0, & \text{c.c.} \end{cases}$

18. $$f(x,y) = \begin{cases} \frac{x-1}{2}, & \text{para } 1 < x < 2 \text{ e } 2 < y < 4 \\ \frac{-x+3}{2}, & \text{para } 2 < x < 3 \text{ e } 2 < y < 4 \\ 0, & \text{fora destes intervalos} \end{cases}$$