UNIVERSIDADE EDUARDO MONDLANE
Faculdade de Ciências
Departamento de Matemática e informática

Probabilidade e Estatística **Resumo teórico** **2º Semestre de 2020**

Tema 1: VARIÁVEIS ALEATÓRIAS E DISTRIBUIÇÕES DE PROBABILIDADE

Variáveis Aleatórias:

Uma variável cujos valores referem-se a eventos aleatórios é chamada **variável aleatória**; seus valores dependem dos resultados de um experimento e a cada valor está associada a uma certa probabilidade. Pode ser **discreta** ou **contínua**, dependendo dos valores que ela assume.

Em geral, denota-se uma variável por letra maiúscula e os valores assumidos por ela por letra minúscula.

As variáveis aleatórias que assumem valores em um conjunto finito ou infinito enumerável são chamadas discretas e aquelas que assumem valores em um intervalo da reta real são chamadas contínuas. Isto é,

Uma variável é dita **Discreta** quando assume valores em pontos isolados ao longo de uma escala (nº finito ou infinito enumerável de valores).**Exemplo:** Nº de alunos na sala

Uma variável é dita **Contínua** quando assume qualquer valor ao longo de um intervalo (nº infinito não enumerável de valores). **Exemplo:** Tempo, temperatura, peso, etc.

Portanto, para atingirmos nosso objetivo, calcular a probabilidade em determinada situação devemos sempre ter em destaque qual a variável aleatória em nosso problema.

Variável Aleatória Discreta:

A variável aleatória é denominada "discreta" quando pode assumir um número finito de valores num intervalo finito. Geralmente seus valores são obtidos através da contagem.

> Uma variável aleatória é discreta quando, entre dois valores seqüenciais não se pode ter um outro valor. As variáveis discretas assumem somente um número finito de valores.

Variável Aleatória Contínua:

A variável aleatória é denominada "contínua" são aquelas que podem assumir infinitos valores num intervalo finito. O conjunto universo da variável contínua possui infinitos elementos.

> A variável aleatória contínua é aquela que se pode atribuir qualquer valor dentro de determinado intervalo

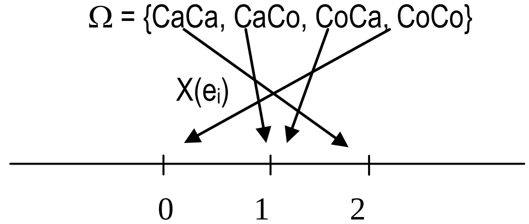
1.1. Variável aleatória discreta

Definição

Muitos experimentos produzem resultados **não numéricos**. Antes de analisá-los é conveniente transformar seus resultados em números. Para isso devemos associar a cada resultado elementar (e_i) do espaço amostral (S) um número real, o que é feito por meio de uma regra ou função denominada **variável aleatória**. Considerando, por exemplo, o experimento *inspecionar um conjunto de caixas conteno 10 peças embaladas manualmente*. Pode-se definir uma variável aleatória como *número de peças boas*. Trata-se de uma variável aleatória discreta, porque ela pode assumir o número finito de valores: 0, 1, 3, ..., 10.

Exemplo 1.1 Como outro exemplo de variável aleatória discreta, consideremos lançamento de duas moedas honesta onde o conceito é ilustrado com um espaço amostral com 4 resultados elementares, ou seja:

Ω = {CaCa, CaCo, CoCa, CoCo}



em que X denota o número de Cara (Ca). Assim definida, X é uma variável aleatória discreta.
Note que para ser discreta, a variável aleatória (v.a.) deve assumir valores em um conjunto finito ou infinito, porém contável. O passo fundamental para entendermos uma v.a. é associar a cada valor de X sua probabilidade, obtendo o que se chama uma **distribuição de probabilidade**.

Exemplo 1.2.: No lançamento de uma moeda 3 vezes e observando a sequência de cara (K) e coroa (C) que aparecem na face de cima.

a) Qual será o espaço amostral desse experimento?.
b) Escreva o subespaço constituido pela ocorrência de duas ou mais caras.

Resolução

a) O espaço amostral consiste em 8 elementos. Para encontrar este número usemos o diagrama de árvore que nos dará o número das possibilidades em três lançamentos da moeda

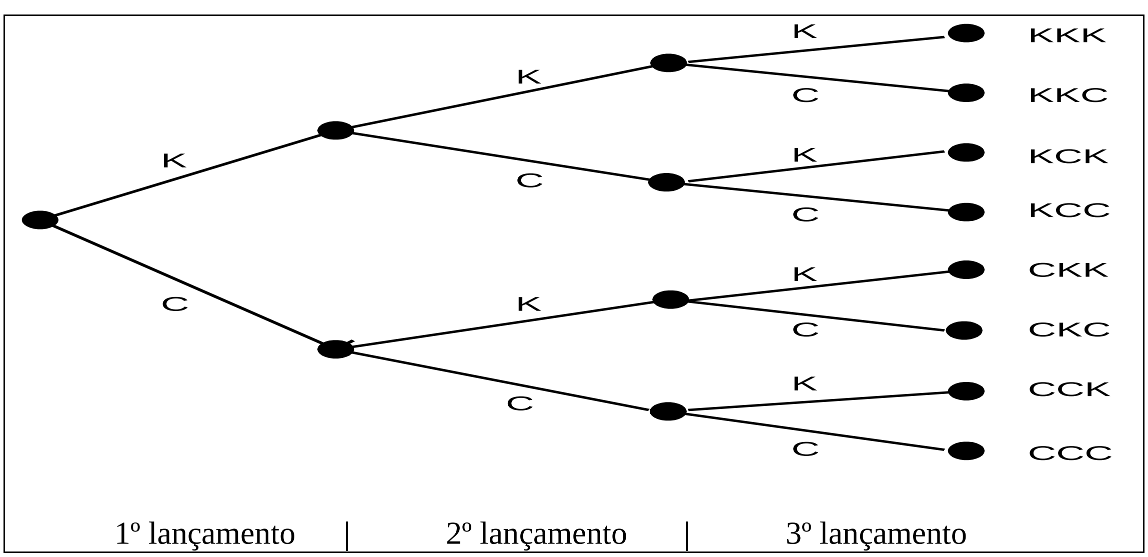


Figura 2.1. Determinação do número de elementos no espaço amostral (teorema de contagem)

O espaço amostral será: S = {kkk, kkc, kck, kcc, ckk, ckc, cck, ccc}
b) O subespaço com duas ou mais caras é : W = {kkk (3), kkc (2), kck(2), ckk(2)}

1.2. Distribuições Discretas de Probabilidade:

Seja X uma variável aleatória discreta e sejam x_1, x_2, ... , x_n os valores de X. A função f(x) é uma distribuição de probabilidade (ou função de probabilidade) se:

(a) $P(X=x) \geq 0, \ \forall x$

(b) $\sum_{i=1}^{n} f(x_i) = 1$

Com base no exemplo anterior, determine a distribuição de probabilidade da variável aleatória X: "Número de caras em dois lances de uma moeda equilibrada". Onde Ca = cara e Co = coroa

S={CoCo, CoCa, CaCo, CaCa}

Eventos (ponto amostral)	CoCo	CaCo ou CoCa	CaCa	Total
$X = x_i$	0	1	2	
$P(X = x_i) = p(x_i)$	1/4	1/2	1/4	1,0

Note que que: $p(x_i)$ é chamada função de probabilidade, que a cada valor de x_i associa sua probabilidade de ocorrência.

A distribuição de probabilidade mostra-nos como a probabilidade total (1,0) é distribuída de acordo com os diferentes valores da variável aleatória. Frequentemente, uma fórmula matemática pode ser usada para representar uma distribuição de probabilidade, em lugar de uma tabela.

Representação gráfica de uma distribuição de probabilidade

Gráfico

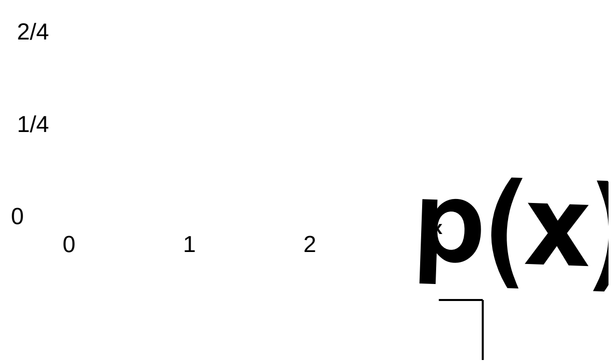


Gráfico para a distribuição dada no Exemplo anterior

Função de Distribuição Acumulada:

A função de distribuição acumulada de uma v.a.d. X é definida por: $F(X) = P(X \leq x) = \sum_{u \leq x} f(u)$

Exemplo 2: Considere o experimento anterior que consiste em lançar duas vezes uma moeda. Seja a variável aleatória X: "número de caras".

A função de distribuição acumulada de X é dada por:

$$F(X) = \begin{cases} 0{,}00;\ se\ x < 0 \\ 0{,}25;\ se\ 0 \le x < 1 \\ 0{,}75:\ se\ 1 \le x < 2 \\ 1{,}00;\ se\ x \ge 2 \end{cases}$$

Parâmetros de uma variável aleatória discreta:

Experança Matemática

Seja X uma variável aleatória discreta que pode assumir os valores x_1, x_2, ... , x_n , com probabilidades $P(x_1)$, $P(x_2)$, ..., $P(x_n)$, respectivamente. Então, a experança de X é dada por: $E(X) = \sum_{i=1}^{n} x_i.P(x_i)$

Propriedades:

1. $E(aX) = aE(X)$
2. $E(X \pm a) = E(X) \pm a$
3. $E(X \pm Y) = E(X) \pm E(Y)$

Variância

Seja X uma variável aleatória discreta com experância finita. Então, a variância de X é dada por:

$Var(X) = E(X^2) - E(X)^2$, onde $E(X^2) = \sum_{i=1}^{n} x_i^2.P(x_i)$

Propriedades:

1. $Var(aX) = a^2 Var(X)$
2. $V(X \pm a) = Var(X)$
3. $Var(X \pm Y) = Var(X) + Var(Y)$, se X e Y são variaveis aleatórias independentes
4. $Var(X \pm Y) = Var(X) + Var(Y) + 2Cov(XY)$, se X e Y são variaveis aleatórias não independentes.

Desvio Padrão:

O desvio padrão de X é dado por $\sigma_X = \sqrt{Var(X)}$

Distribuição de probabilidade conjunta de duas variáveis aleatórias:

$$P_{ij} = P(X = i \wedge Y = j).$$

Costuma-se indicar com uma tabela de dupla entrada.

Por exemplo

X \ Y	2	4
1	$1/4$	$1/3$
2	$1/12$	$1/6$
3	$1/12$	$1/12$

Assim, $P(X=1 \wedge Y=2)=\frac{1}{4}$ e $P(X=3 \wedge Y=4)=\frac{1}{12}$. Note-se que

$X \backslash Y$	2	4	$P(X=i)$
1	$1/4$	$1/3$	$7/12$
2	$1/12$	$1/6$	$3/12$
3	$1/12$	$1/12$	$2/12$
$P(Y=i)$	$5/12$	$7/12$	1

Portanto a função de probabilidade X , também dita função de probabilidade marginal de X , é

$X=i$	1	2	3
$P(X=i)$	$7/12$	$3/12$	$2/12$

e a função de probabilidade marginal de Y é

$Y=j$	2	4
$P(Y=j)$	$5/12$	$7/12$

Também se podem definir, por exemplo, a função da probabilidade de X condicionada por $Y=2: X/Y=2$:

$(X/Y=2)=i$	1	2	3
$P((X/Y=2)=i)$	$\frac{1/4}{5/12}$	$\frac{1/12}{5/12}$	$\frac{1/12}{5/12}$

ou por exemplo, a função de probabilidade Y condicionada por $X=3: Y/X=3$:

$(Y/X=3)=j$	2	4
$P((X/Y=3)=j)$	$\frac{1/12}{2/12}$	$\frac{1/12}{2/12}$

Se, ainda, fizermos por exemplo $Z=X+Y$ a sua função de probabilidade é

$Z=i$	3	4	5	6	7
$P(Z=i)$	$1/4$	$1/12$	$5/12$	$1/6$	$1/12$

Covariância é uma medida de associação (relação) linear entre duas variáveis aleatórias. Se X e Y são duas v.a., a covariância entre elas é definida por:

$$Cov(X,Y)=E(XY)-E(X).E(Y)$$

1.3. Distribuições Contínuas de Probabilidade:

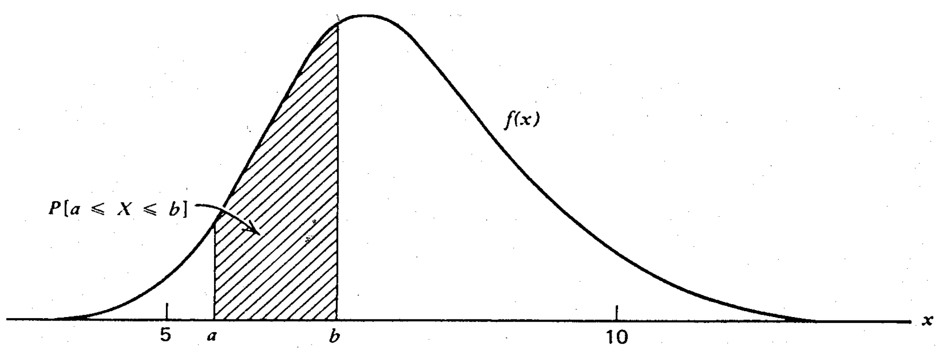
Seja X uma variável aleatória contínua. A função f(x) é uma função densidade de probabilidade se:

a) $f(x) \geq 0$, (não negativa)

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

c) P(a <X< b)= $\int_a^b f(x)dx$, área sob a curva entre os pontos a e b;

d) $P(X = x) = 0$



Exemplo 3: Considere a função densidade dada abaixo.

$$f(x) = \begin{cases} \frac{1}{9}x^2, \text{se } 0 < x < 3 \\ 0, \textit{caso contrário} \end{cases}$$

Função de Distribuição Acumulada:

A função de distribuição acumulada de uma v.a.c. X é definida por

$$F(X)=P(X \leq x) = \int_{-\infty}^{x} f(u)du$$

Considerando a função densidade anterior, a função de distribuição de X é dada por:

$$F(x) = \begin{cases} 0; \text{se x} \leq 0 \\ \frac{1}{27}x^3; \text{se } 0 < \text{x} < 3 \\ 1; \text{se x} \geq 3 \end{cases}$$

OBS: $f(x) = \frac{dF(x)}{dx}$

Determine a probabilidade de da variável aleatoria X:

a) $P(x < 1)$

b) $P(x < 2)$

c) $P(1 < x < 2)$

Solução:

a) $P(x<1)=F(1)=\frac{1}{27}1^3=\frac{1}{27}$

b) $P(x<2)=F(2)=\frac{1}{27}2^3=\frac{8}{27}$

c) $P(1<x<2)=F(2)-F(1)=\frac{1}{27}2^3-\frac{1}{27}1^3=\frac{8}{27}-\frac{1}{27}=\frac{7}{27}$

Exemplo 4: Suponha-se que X seja uma VAC com fdp dada por: $f(x)=\begin{cases}2x, \text{ se } 0<\text{x}<1\\ 0, \text{caso contrário}\end{cases}$

Determine a função de distribuição acumulada.

$$\text{F(X)=P(X} \le \text{x)} = \int_{-\infty}^{x} f(u)du = \int_{-\infty}^{x} 2udu = \begin{cases}0\ , se\ x\le 0\\ x^2,\ \ se\ \ 0<x\le 1\\ 1,\ \ se\ x>1\end{cases}$$

Exemplo 5: Seja X a corrente de um fio de cobre. A faixa de valores de X é [0 á 20mA]. A função densidade de probabilidade de X é: f (x) = 0,05 para $0\le x\le 20mA$.

a) Qual a probabilidade da corrente ser menor que 10 miliampéres?
b) Qual a probabilidade da corrente variar entre 5 e 20 miliampéres?

Solução:
Note que f (x) é zero para os valores que ela não foi definida.

a) $P(x<10)=\int_0^{10} f(x)dx=\int_0^{10} 0{,}05dx=0{,}05x\,|^{10}\ -\ 0{,}05x\,|^{0}=0{,}5-0=0{,}5$

b) $P(5<x<20)=\int_5^{20} f(x)dx=\int_5^{20} 0{,}05dx=0{,}05x\,|^{20}\ -0{,}05x\,|^{5}=1-0{,}25=0{,}75$

Parâmetros de uma variável aleatória contínua:

Experança:

Seja X uma variável aleatória contínua com função densidade de probabilidade f(x). Então, a expectância de X é dada por

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

Exemplo 5: Para a função densidade dada anteriormente, $E(X)=\int_0^3 x\frac{1}{9}x^2dx=\frac{1}{36}x^4\Big]_0^3=2{,}25$

Variância:
Seja X uma variável aleatória contínua com função densidade f(x). Então, a variância de X é dada por:

$$Var(X)=E(X^2)-[E(X)]^2 \text{, onde } E(X^2)=\int_{-\infty}^{+\infty} x^2 f(x)dx$$

Desvio Padrão: O desvio padrão de X é dado por: $\sigma_X=\sqrt{Var(X)}$

Exemplo 6: Considerando a função densidade dada por

$$f(x) = \begin{cases} \frac{1}{9}x^2, \text{se } 0 < x < 3 \\ 0, \text{caso contrário} \end{cases}$$

$$E(X^2) = \int_0^3 x^2 \frac{1}{9} x^2 . dx = \frac{1}{45} x^5 \Big]_0^3 = 5{,}4$$

Então,

$Var(X) = 5{,}40 - (2{,}25)^2 = 0{,}3375$ e

$$\sigma_x = \sqrt{0{,}3375} \cong 0{,}5809$$

2. Principais distribuições discretas de probabilidade

2.1 Introdução

Existem muitas relações, tais como a exponencial, a quadrática, a linear, a parabólica, que são usadas para descrever a funcionalidade entre as variáveis determinísticas. No caso das variáveis aleatórias, certas distribuições de probabilidade são usadas com mais frequência para descrever a grande maioria dos fenômenos físicos.

As distribuições de probabilidade, para variáveis aleatórias discretas, mais comumente usadas e que desenvolvem um papel importante na metodologia estatística são: Bernoulli, Binomial, Poisson, Geométrica, Pascal, Hipergeométrica e Multinomial.

.

2.2. Processo de Bernoulli

Esse processo é análogo àquele de jogar uma moeda. As seguintes suposições se aplicam:

a) Cada experimento é dito ser uma tentativa. Existe uma série de tentativas, cada uma tendo dois resultados: sucesso ou falha;
b) A probabilidade de sucesso é igual a algum valor constante para todas as tentativas;
c) Os resultados sucessivos são estatisticamente independentes. A probabilidade de sucesso na próxima tentativa não pode variar, não importando quantos sucessos ou falhas tenham sido obtidos.

O processo de Bernoulli é comumente utilizado em aplicações de engenharia envolvendo controle de qualidade. Cada novo item criado no processo de produção pode ser considerado como uma tentativa resultando em uma unidade com ou sem defeito. Esse processo não se limita a objetos; podendo ser usado em pesquisas eleitorais e de preferências dos consumidores por determinados produtos.

A propriedade desse modelo é que a variável aleatória X é baseada somente em dois resultados possíveis: sucesso (**x = 1**) e o fracasso (**x = 0**),mesmo sendo considerada uma experiência aleatória. Em Bernoulli, o somatório entre p e q sempre será igual à 1: $p + q = 1$

A função da **Distribuição de Bernoulli** é dada por: $P(X = x) = p^x \times q^{1-x}$ em que:

x = nº de sucesso em uma única tentativa de experimentos (só assumirá o valor de 1 ou 0);
p = probabilidade de sucesso; **q** = probabilidade de fracasso

Exemplo 2.1: Qual é a probabilidade de obter face 5 em uma única tentativa do arremesso de um dado?

Sabe-se que **x** = 1 **p** = 1/6 **q** = 5/6

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

2.3 Distribuição Binomial

A distribuição binomial é aplicada frequentemente para descrever controle estatístico de qualidade de uma população. Tem-se interesse principalmente em duas categorias: item defeituoso ou insatisfatório *versus* item bom ou satisfatório ou sucesso e falha que tenham ocorrido em uma amostra de tamanho fixo. A distribuição binomial é aplicada a eventos provenientes de uma série de experimentos aleatórios, que constituem o chamado Processo de Bernoulli.

A distribuição binomial está relacionada ao número de sucessos encontrados em um número específico de tentativas de um processo de Bernoulli. Ela é caracterizada por dois parâmetros:
p ($0 \le p \le 1$), a probabilidade de sucesso, e ***n*** o número de tentativas (número positivo).

Num exemplo das moedas (exemplo 1.2), dado anteriormente, construiu-se uma árvore de probabilidade, onde 3 tentativas (lançamentos) foram feitos com o objectivo de obter face "Cara" ou K. Imagine agora a enorme árvore (aproximadamente 6 metros de altura) que se teria que construir se fossem feitas 20 tentativas! De modo a evitar isto, a equação abaixo é usada para descrever a probabilidade de uma variável aleatória *X*, que denota o número de sucessos em *n* tentativas do processo de Bernoulli, utilizando *k* itens.

$X \sim bin\,(n, p)$ $\quad P(X=x)=\binom{n}{x}p^{x}(1-p)^{n-x}$ para x = 0,1,2,...,*n* sendo $\binom{n}{x}=\frac{n!}{x!(n-x)!}$

Exemplo 2.2. Calcular as probabilidades de um casal ter meninas em três nascimentos.

Solução:

a) *x* = 0; *p* = 0,48; *n* =3

$$P(X=0)=\frac{3\times 2\times 1}{1\times(3-0)!}\times 0{,}48^{0}\times(1-0{,}48)^{3-0}=0{,}14$$

b) *x*=1; *P* = 0,48; *n*=3

$$P(X=1)=\frac{3\times 2\times 1}{1\times(3-1)!}\times 0{,}48\times(1-0{,}48)^{3-1}=0{,}39$$

c) *x*=2; *p*=0.48; *n*=3

$$P(X=2)=\frac{3\times 2\times 1}{2\times 1\times(3-2)!}\times 0.48^{2}\times(1-0.48)^{3-2}=0.36$$

d) *x*=3; *p*=0.48; *n*=3

$$P(X=3)=\frac{3\times 2\times 1}{3\times 2\times 1\times(3-3)!}\times 0.48^{3}\times(1-0.48)^{3-3}=0.11$$

Nas Figuras abaixo são apresentados 3 distribuições binomiais, onde se observa que à medida que o número de tentativas cresce, as probabilidades vão diminuindo e ficando menos concentradas, tendo a distribuição um formato de sino, independente do parâmetro π, a probabilidade de sucesso.

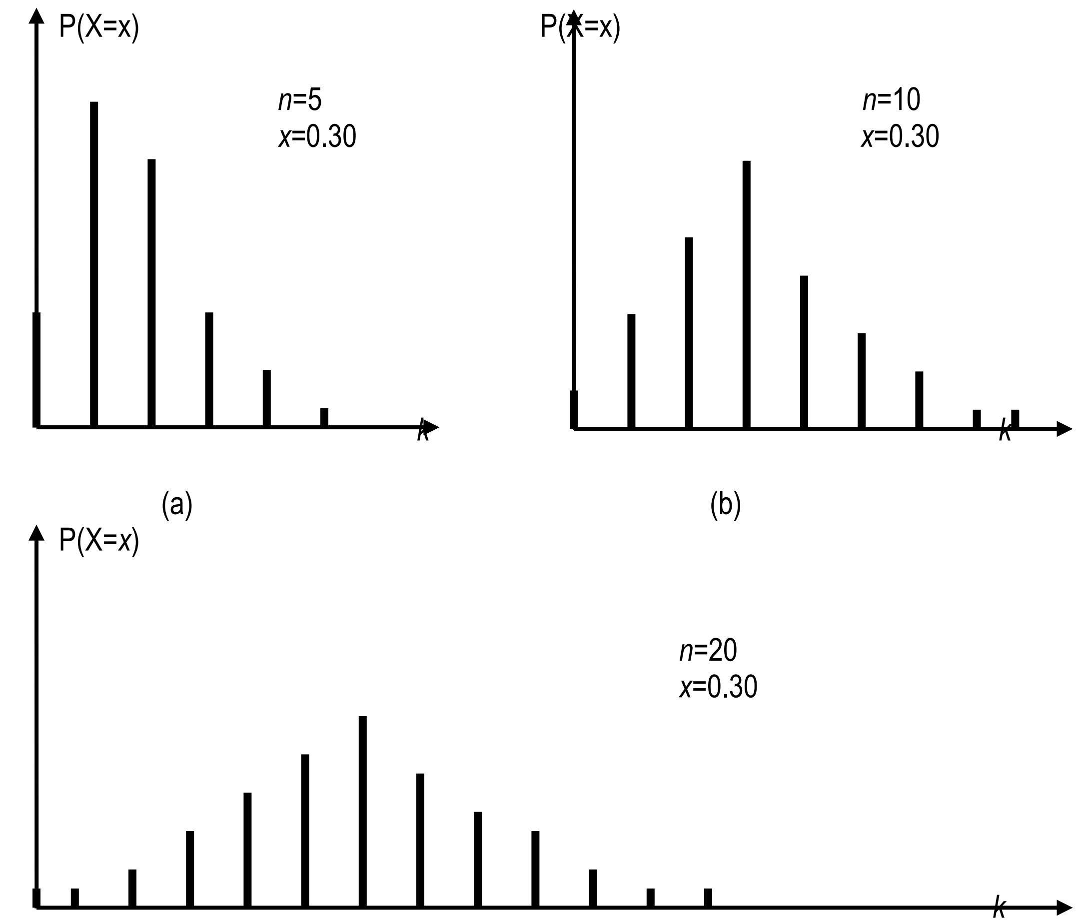


Figura 1 – Distribuições Binomiais

Valor Esperado e Variância para distribuição Binomial

O valor esperado e a variância da variável aleatória X, quantos sucessos obtidos em um certo número de tentativas, são dados, respectivamente, por:

$$E(X) = np \quad \text{e} \quad Var(X) = np(1-p)$$

Exemplo 2.3: Três em cada quatro alunos de uma universidade fizeram curso de curta duração antes de realizar exame de admissão para o ingresso a referida universidade. Se 16 alunos são selecionados ao acaso, qual é a probabilidade de que:

a) Pelo menos 12 tenham feito um curso de curta duração?

Seja X o número de alunos que fizeram o curso de curta duração

p: probabilidade de um aluno, selecionado ao acaso, ter feito o curso de curta duração; p = 0,75.

X ~bin (16; 0,75), $P(X=x)=\binom{n}{x}p^x(1-p)^{n-x} \Rightarrow P(X=x)=\binom{16}{x}\times 0{,}75^x(1-0{,}75)^{16-x}$

ou seja, a variável aleatória X tem distribuição binomial com parâmetros n = 16 e p = 0,75.

Assim, a probabilidade de que pelo menos 12 tenham feito curso de curta duração é dada por:

P(X ≥ 12) = P(X=12) + P(X=13) + P(X=14) + P(X=15) + P(X=16) = 0,2252+0,2079+0,1336+ +0,0535+0,0100= 0,6302

b) No máximo 13 tenham feito curso de curta duração?

Utilizando a função de distribuição apresentada no item (a) temos,

P(X ≤ 13) = P(X=0) + P(X =1) + … + P(X=13) = 0,0000 + … + 0,2079 = 0,8029
ou
P(X ≤ 13) = 1 - P(X ≥ 14) = 1 – (P(X =14) + P(X =15) + P(X =16) = 0,8029

c) Exactamente 12 tenham feito curso de curta duração?

Utilizando a função de probabilidade apresentada no item (a) temos, P(X =12) = 0,2252

d) Em um grupo de 80 alunos selecionados ao acaso, qual é o número esperado de alunos que fizeram curso de curta duração? E a variância?

Y: número de alunos que fizeram curso de curta duração entre os 80 selecionados
Y~Bin (80 ; 0,75)
O número esperado de alunos que fizeram curso de curta duração é dado por:

μ = E(X) = np = 80 x 0,75 = **60**

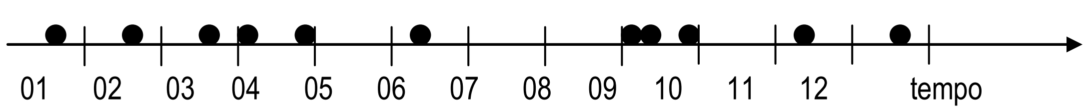
A variância é dada por: σ^2 = Var(x) = n x p x (1-p) =80 x 0,75 x 0,25 = 15

2.3. Distribuição de Poisson

A distribuição de Poisson é uma das mais usadas para variáveis aleatórias discretas. Sua aplicação mais comum é na descrição de dados sobre, por exemplo, número diário de telefonemas a uma central telefônica, número de carros que passam por um cruzamento (ou uma cabine de pedágio) durante um certo período de tempo e em análises de confiança em uma linha de produção (saber probabilidades de falhas). Os eventos devem ocorrer em um certo intervalo de tempo ou espaço.

Processo de Poisson

Imagine que se está interessado em saber o número de lâmpadas queimadas em uma rua, durante 12 dias do mês de agosto. Cada lâmpada queimada é considerada um evento. Obtém-se o seguinte resultado, tendo-se uma taxa média de λ = 1 (uma lâmpada queimada diariamente).



Todo o processo de Poisson tem as seguintes propriedades:

a) O número de eventos ocorrendo em um segmentoou intervalo de tempo ou espaço é independente do número de eventos ocorridos no intervalo anterior; o processo de Poisson não tem memória;
b) A taxa média do processo, λ, deve permanecer constante durante o período de tempo e espaço considerados;
c) Quanto menor o segmento (intervalo) de tempo ou de espaço, menor a probabilidade de ocorrer mais de um evento naquele segmento. A probabilidade de ocorrência de 2 ou mais eventos se aproxima de zero, quando o tamanho do intervalo se aproxima de zero.

A distribuição de Poisson é dada por: $P(X = x) = \dfrac{(\lambda)^x e^{-\lambda}}{x!}$ para x = 0,1,2,...

em que o parâmteros λ representa a taxa média do processo em um intervalo de tempo/espaço.

Exemplo 2.4: Considere o horário de pico do uso da portagem (saída para feriadão, por exemplo). Suponha que 600 carros por hora passem pela portagem. Está-se interessado no número de carros que chegam, variável X, durante um período de 12 segundos. Assim, tem-se λ = 600. As probabilidades de X são dadas abaixo:

$$P(X = 0) = \frac{2^0 e^{-2}}{0!} = 0.1353$$

$$P(X = 1) == \frac{2^1 e^{-2}}{1!} = 0.2707$$

$$P(X = 2) = \frac{2^2 e^{-2}}{2!} = 0.2707$$

$$P(X = 3) = \frac{2^3 e^{-2}}{3!} = 0.0361$$

$$\vdots$$

$$P(X = 10) = \frac{2^{10} e^{-2}}{10!} = 0.0000$$

Observa-se que à medida que X aumenta, as probabilidades de Poisson tendem rapidamente a zero. Os resultados anteriores mudam completamente se os valores dos parâmetros λ e o intervalo de tempo/espaço forem alterados. Embora 2 ou mais carros possam chegar à ponte em um intervalo de tempo muito curto, a probabilidade deles chegarem ao mesmo tempo é considerada zero. Mesmo havendo alguns exemplos práticos em que os carros cheguem ao mesmo tempo, a distribuição de Poisson é um modelo teórico que satisfaz plenamente.

Exemplo 2.5. Supondo que o número médio de bactérias por litro de água purificada é 2, qual é a probabilidade que 5 ou mais bactérias sejam encontradas em uma amostra de 3 litros de água?

Sendo λ = 2.3 = 6, o número médio de bactérias em 3 litros de água, então:

$$P(x \geq 5) = 1 - P(X \leq 4) = 1 - \sum_{x=0}^{4} \frac{e^{-6} 6^x}{x!} = 1 - 0{,}2851 = 0{,}7149$$

Exemplo 2.6. Em uma população, seja X o número de descendentes produzidos por família/geração. Assumindo que $\overline{X} = \lambda = 2$, qual a probabilidade de famílias selecionada ter $X = 4$ descendentes?

$$P(X = 4) = \frac{e^{-2} \times 2^4}{4!} = 0{,}0902$$

Valor Esperado e Variância para distribuição de Poisson

O valor esperado e a variância de uma distribuição de Poisson são iguais a:

$E(X) = \lambda$ e $Var(X) = \lambda$

Assim, vê-se que o valor esperado de carros passando pela portagem deve dobrar se dobramos o tempo de observação, o mesmo acontecendo com a variância.

2.4. Distribuição geométrica

A distribuição geométrica é aplicada quando está interessado na probabilidade de acontecer na 1º vez algum fenômeno acerca algum número de tentativas. Ela é considerada um caso particular da Distribuição de Pascal pois, enquanto a função geométrica funciona apenas quando se é desejado a 1º vez do acontecimento, a distribuição de pascal preocupa-se quantas vezes o experimento será repetido para que seja um sucesso na k-ésima vez. Em suma, a distribuição geométrica é uma derivação da função da distribuição de pascal. o somatório entre p e q sempre será igual à 1.

$$p+q=1$$

A distribuição geométrica é dada por:

$$P(X=x)=p(1-p)^{x-1}=q^{x-1}p \quad \text{onde a} \quad E(X)=\frac{1}{p} \quad \text{e} \quad Var(X)=\frac{(1-p)}{p^2}$$

Em que: **x** = número de tentativas necessária para o aparecimento do primeiro sucesso;
p = probabilidade de sucesso; **q** = probabilidade de fracasso.

Exemplo 2.5: Uma urna tem 10 bolas brancas e 40 bolas pretas. Qual a probabilidade de que a 6ª bola retirada com reposição seja a 1ª bola branca?

x = número de retiradas necessárias para obter a primeira bola branca; **p** = 10/50 = 0,2 **q** = 40/50 = 0,8

$$P(X=6)=p(1-p)^{x-1}=(0{,}8)^{6-1}\times 0{,}2=0{,}8^5\times 0{,}2=0{,}065536$$

2.5. Distribuição de Pascal ou Binomial Negativa

A distribuição de Pascal envolve quantas repetições são necessárias para que ocorra, em um experimento, um evento desejado (sucesso) pela k-enésima vez. Ao contrário da distribuição geométrica que considera **k** = 1, sendo ele o número desejado para que se ocorra a chance, nesse caso, pela 1º vez. Dessa forma, a distribuição de pascal pode ser utilizada para calcular probabilidades feitas pela distribuição geométrica por ser considerada sua equação geral. A somatória entre p e q sempre será igual à 1: $p+q=1$

A distribuição de Pascal é dada por:

$$P(X=x)=\binom{x-1}{r-1}p^r\times q^{x-r} \qquad \text{Lembre – se que } \binom{n}{x}=\frac{n!}{(n-x)x!}$$

Em que:

x = número de recorrência exigido para que um evento aconteça pela k-enésima vez;
p = probabilidade de sucesso; **q** = probabilidade de fracasso; **r** = número de sucessos na amostra.

$$E(X)=\frac{r}{p} \quad \text{e} \quad Var(X)=\frac{r(1-p)}{p^2}$$

Exemplo 2.6: A probabilidade de que um sinal de trânsito esteja aberto numa esquina é 0,20. Qual a probabilidade de que seja necessário passar pelo local 10 vezes para encontrá-lo aberto pela 4ª vez?

x = número de passagens pela esquina. **r** = 4 **p** = 0,20 **q** = 0,80

$$P(X = 10) = \binom{10-1}{4-1} \times 0{,}2^4 \times 0{,}8^{10-4} = \binom{9}{3} \times 0{,}2^4 \times 0{,}8^6 = 0{,}035232$$

2.6. Distribuição Hipergeométrica

Essa distribuição é uma distribuição de probabilidade discreta que é adequada quando consideramos extrações casuais feitas sem reposição de uma população dividida segundo dois atributos.

Para ilustrar, considere uma população de N objetos, r dos quais têm o atributo A e (N - r) têm o atributo B. Um grupo de n elementos é escolhido ao acaso, sem reposição. Estamos interessados em calcular a probabilidade de que esse grupo contenha x elementos com o atributo A. portanto, diremos que uma variável aleatória X tem distribuição hipergeométrica de parâmetros, $N, r \ e \ n$ se sua função de probabilidade for dada da maneira abaixo. Denotamos $X \sim Hgeo(N, r, n)$:

$$P(X = x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} \qquad \text{lembre-se que} \qquad \binom{n}{x} = \frac{n!}{(n-x)x!}$$

Em que:

x = característica escolhida (Número de sucesso que se pretende);
N = número de itens no experimento;
n= número de itens na amostra, pode ser uma estimação;
r = número de sucessos na amostra.

A distribuição hipergeométrica é aplicada quando o objectivo do cálculo da probabilidade de um evento envolva o número de sucessos em uma amostra selecionada aleatoriamente, ou seja, a saída do elemento com a característica envolvida e almejada.

$$E(X) = np \qquad \text{e} \qquad Var(X) = np(1-p)\frac{N-n}{N-1}$$

Exemplo 2.7: Pequenos motores são guardados em caixas de 50 unidades. Um inspetor de qualidade examina cada caixa, antes da posterior remessa, testando 5 motores. Se nenhum motor for defeituoso, a caixa é aceite. Se pelo menos um motor for defeituoso, todos os 50 motores são testados. Há 6 motores defeituosos numa caixa. Qual a probabilidade de que seja necessário examinar todos os motores dessa caixa?

*Pelo menos um implica que se for encontrado um motor defeituoso, ou dois motores defeituosos,..., ou seis motores defeituosos, os 50 motores serão testados, assim, os 50 motores não serão testados se e somente se nenhum motor entre os 6 for defeituoso. Logo basta calcular **P(x**=0) e subtrair de 1.*

x = número de motores defeituosos da amostra. **N** = 50 **r** = 6 **n** = 5

$$P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0) = 1 - \frac{\binom{6}{0}\binom{44}{5}}{\binom{50}{5}} = 1 - 0{,}512567728 = 0{,}487432271$$

2.7. Distribuição Multinomial

É uma extensão da distribuição binomial para mais de dois resultados possíveis, ou seja, uma generalização dessa distribuição. Por exemplo, quando se quer descobrir a probabilidade da ocorrência de eventos de um lançamento de dados utiliza-se a distribuição multinomial por se tratar de mais de dois resultados possíveis. o somatório entre p e q sempre será igual à 1.

$$p + q = 1$$

A distribuição multinomial é dada por:

$$P(X_1 = n_1, X_2 = n_2, ..., X_k = n_k) = \frac{n!}{n_1!n_2!...n_k!} \times p_1^{n_1} \times p_2^{n_2} \times ... \times p_k^{n_k}$$

Em que: **x** = número de sucessos em n tentativas;
p = probabilidade de sucesso;
q = probabilidade de fracasso;
k = número de sucessos na amostra.

Exemplo 2.8: Uma urna tem 6 bolas brancas, 4 pretas e 5 azuis. Retiram-se 8 bolas com reposição. Qual a probabilidade de sair 4 bolas brancas, 2 pretas e 2 azuis?

Resolução:

$P_1 = P(B) = 6/15 = 2/5$ $P_2 = P(P) = 4/15$ $P_3 = P(A) = 5/15 = 1/3$

X_1 = saída de 4 bolas brancas X_2 = saída de 2 bolas pretas X_3 = saída de 2 bolas azuis

Aplicando a fórmula:

$P(X_1 = 4, X_2 = 2, X_3 = 2) = = \frac{8!}{4!2!\,2!} \times \left(\frac{2}{5}\right)^4 \times \left(\frac{4}{15}\right)^2 \times \left(\frac{1}{3}\right)^2 = 0{,}084954074$

Ficha de Exercicios Nº 1

1. Determine se é dada uma distribuição de probabilidade. Nos casos em que não é descrita uma distribuição de probabilidade, identifique a condição que não é satisfeita. E quando for descrita uma ***distribuição de probabilidade***, determine sua média, variância e desvio-padrão.

a)

X	0	1	2	3
P(X)	0,2	0,3	0,3	0,2

b)

X	-2	-1	0
P(X)	0,25	0,50	0,20

c)

X	4	9	20
P(X)	-0,3	1,0	0,3

d)

X	2	3	5	6
P(X)	0,15	0,20	0,40	0,35

2. Consideremos a ocorrência de acidentes de trabalho em uma empresa, durante os 6 primeiros meses do ano. Determine a probabilidade da ocorrência de acidentes para cada mês.

x_i	Janeiro	Fevereiro	Março	Abril	Maio	Junho
f_i	2	3	5	4	6	5

3. Elabore a distribuição da probabilidade do número de "caras" obtido ao se lançar uma moeda 4 vezes.

4. Lança-se uma moeda 3 vezes, considerando que a moeda seja honesta, [*P(cara)* = *P(coroa)*].
   a) Determine o espaço amostral.
   b) Determinar a tabela de distribuição de probabilidades de X (número de caras).
   c) O número esperado de defeitos.
   d) variancia e desvio padrão.

5. Numa loja de electrodoméstico a procura diária de aspiradores é uma variável aleatória discreta X com função de probabilidade:

x	0	1	2	3	4
P(x)	0,2	b	0,3	0,2	0,1

Determine o valor do b;

6. Qual a esperança matemática e o desvio-padrão de um jogo no qual se pode ganhar \$50,00 com probabilidade 0,2; \$30,00 com probabilidade 0,3 e \$10,00 com probabilidade 0,5?

7. Seja X: v.a.contínua com função densidade de probabilidade (f.d.p.) dada por :

$$f(x) = \begin{cases} 0 \quad \text{se} \quad x < 0 \ \text{ ou } \ x > 2 \\ \dfrac{x}{2} \quad \text{se} \quad 0 \le x \le 2 \end{cases}$$

Determinar :
   a) $P(0,5 < x \le 1,0)$
   b) $P(1,0 < x \le 1,5)$
   c) E(x) e V(x).

8. Observe a função de distribuição acumulada *F*(x) abaixo e calcule as probabilidades de:

$$F(x) = \begin{cases} 0, & \text{se} \quad x < 0 \\ \dfrac{x^2}{50}, & \text{se} \quad 0 \le x < 5 \\ -\dfrac{x^2}{50} + \dfrac{2}{5}x - 1, & \text{se} \quad 5 \le x < 10 \\ 1, & \text{se} \quad x \ge 10 \end{cases}$$

   a) $P(x \le 2)$
   b) $P(3 \le x \le 8)$
   c) $P(x \le 6)$

9. Suponha que a probabilidade dos pais terem um filho(a) com cabelos loiros seja ¼. Se houverem 4 crianças na família, qual é a probabilidade de que metade delas terem cabelos loiros?

10. Estima-se que cerca de 30% dos frangos congelados contenham suficiente número de bactérias *salmonelas* causadoras de doenças, se forem assados inadequadamente. Um consumidor compra 12 frangos congelados. Qual é a probabilidade do consumidor ter mais de 6 frangos contaminados?

11. Seja bin (n,p). Sabendo-se que E(X) = 12 e V(X) = 4. Determinar: n, p e q.

12. Um departamento de polícia recebe em média 5 solicitações por hora. Qual a probabilidade de receber duas solicitações numa hora selecionada aleatoriamente?

13. Suponha que o número de avarias numa central de telecomunicação siga uma distribuição de *Poisson* com $\lambda = 2$. Então a probabilidade da central apresentar mais de 2 avarias será?

14. A probabilidade de um alinhamento ótico bem sucedido na montagem de produto de armazenamento de dados é de 0.80. Assuma que as tentativas são independentes. Qual é a probabilidade de que o primeiro alinhamento bem sucedido requeira exatamente quatro tentativas?

15. Na Tmcel, uma empresa fornecedora de serviços de telefonia movel, apenas 35% das chamadas são relacionadas a reclemações sobre erros nas facturas emitidas pela empresa. Qual é a probabilidade de ocorrer a segunda reclamação da conta na sétima chamada?

Links para assistir alguns videos:

1. https://www.youtube.com/watch?v=gxqFgxp0hCA (variavel aleatoria)
2. **https://www.youtube.com/watch?v=tTTPctReBeM** (variavel aleatoria discreta)
3. **https://www.youtube.com/watch?v=C0G1g77l05s&t=602s** (variavel aleatoria continua)
4. **https://www.youtube.com/watch?v=A6W994ZIzpI** **(binomial)**
5. **https://www.youtube.com/watch?v=OIOtwRW-Mc8** **(poisson)**
6. **https://www.youtube.com/watch?v=RcCae3PDZuw** **(Geometrica)**
7. **https://www.youtube.com/watch?v=iUdTi-xOcVQ** **(binomial negativa ou pascal)**
8. **https://www.youtube.com/watch?v=vF7xbXRIdc0** **(hipergeometrica)**