Distribuições de Probabilidades

Quando aplicamos a Estatística na resolução de problemas administrativos, verificamos que muitos problemas apresentam as mesmas características o que nos permite estabelecer um modelo teórico para determinação da solução de problemas.

Os componentes principais de um modelo estatístico teórico:

1. Os possíveis valores que a variável aleatória X pode assumir;
2. A função de probabilidade associada à variável aleatória X;
3. O valor esperado da variável aleatória X;
4. A variância e o desvio-padrão da variável aleatória X.

Há dois tipos de distribuições teóricas que correspondem a diferentes tipos de dados ou variáveis aleatórias: a distribuição discreta e a distribuição contínua.

Distribuições Discretas

Descreve quantidades aleatórias (dados de interesse) que podem assumir valores particulares e os valores são **finitos.** Por exemplo, uma *variável aleatória discreta* pode assumir somente os valores 0 e 1, ou qualquer inteiro não negativo, etc. Um exemplo de variável climatológica discreta são as tempestades com granizo.

Distribuição de Bernoulli

Característica do modelo

Se uma variável aleatória X **só pode** assumir os valores **0** (fracasso) e **1** (sucesso) com $P(X = 0) = q$ e $P(X = 1) = p$ com $p + q = 1$, então diremos que a variável aleatória X admite distribuição de Bernoulli.

Discrição do modelo

1. $X = \{0,1\}$
2. $P(X = 0) = q$ e $P(X = 1) = p$;
3. $E(X) = p$;
4. $\sigma^2 = Var(X) = p \times q$ e $\sigma = Dp(X) = \sqrt{p \times q}$

Podemos escrever o modelo do seguinte modo:

$$P(X = x) = p^x \cdot q^{1-x}$$

onde $q = 1 - p$.

- Esperança (média) e Variância:

Calcularemos a média e a variância da variável com distribuição de Bernoulli assim:

X	P(X)	X . P(X)	X² . P(X)
0	q	0	0
1	p	p	p
	1	p	p

$E(X) = p$ e $Var(X) = p - p^2 = p(1 - p) = p \cdot q$

EXEMPLO:

No lançamento de uma moeda, a variável aleatória **X** denota o número de *caras* obtidas.

1. $X = \{0,1\}$;
2. $P(X = 0) = 1/2$ e $P(X = 1) = 1/2$;
3. $E(X) = 0 \times 1/2 + 1 \times 1/2 = 1/2$;

4. $\sigma^2 = Var(X) = 1/2 \times 1/2 = ¼$ e $\sigma = Dp(X) = \sqrt{\frac{1}{4}} = 1/2$.

EXERCÍCIO:

Uma urna contém 20 bolas brancas e 30 bolas vermelhas. Uma bola é retirada da urna e a variável aleatória X denota o número de bolas vermelhas obtidas. Calcule a média E(X), a Var(X) e o desvio-padrão de X.

Solução:

Temos: $X = \begin{cases} 0 & \rightarrow \quad q = 20/50 = 2/5 \\ 1 & \rightarrow \quad p = 30/50 = 3/5 \end{cases}$ $\therefore$ $P(X=x) = (2/5)^{x} \cdot (3/5)^{1-x}$

$E(X) = p\ 2/5$ $Var(X) = p.q = (2/5).(3/5) = 6/25$

Distribuição Binomial

1. CONCEITUAÇÃO

Vamos, neste item, considerar experimentos que satisfaçam as seguintes condições:

a. O experimento deve ser repetido, nas mesmas condições, um número finito de vezes (**n**).

b. As provas repetidas devem ser independentes, isto é, o resultado de uma não deve afetar os resultados das sucessivas.

c. Em cada prova deve aparecer um dos dois possíveis resultados: **sucesso** e **insucesso.**

d. No decorrer do experimento, a probabilidade **p** do sucesso e a probabilidade **q** (q = 1 – p) do insucesso manter-se-ão constantes.

Resolveremos problemas do tipo: determinar a probabilidade de se obterem **k** *sucessos* em **n** tentativas.

O experimento "*obtenção de caras em cinco lançamentos sucessivos e independentes de uma moeda*" satisfaz essas condições.

Sabemos que, quando da realização de um experimento qualquer em uma única tentativa, se a probabilidade de realização de um evento (sucesso) é p, a probabilidade de não-realização desse mesmo evento (insucesso) é **1 – p = q.**

Suponhamos, agora, que realizemos a mesma prova n vezes sucessivas e independentes. A probabilidade de que um evento se realize **k** vezes nas provas é dada pela função:

$$\boxed{f(x) = P(X = k) = \binom{n}{k} p^k q^{n-k}}$$

Na qual:

P(X = k) é a probabilidade de que o evento se realize **k** vezes em **n** provas;

p é a probabilidade de que o evento se realize em uma só prova – **sucesso**;

q é a probabilidade de que o evento não se realize no decurso dessa prova – **insucesso**;

$\binom{n}{k}$ é o coeficiente binomial de n sobre k, igual a $\frac{n!}{k!(n-k)!}$

Essa função, denominada **lei binomial**, define a distribuição **binomial.**

EXERCÍCIOS RESOLVIDOS

1. Uma moeda é lançada 5 vezes seguidas e independentes. Calcule a probabilidade de serem obtidas 3 caras nessas 5 provas?

Solução:

Temos:

N = 5 e k = 3

Pela lei binomial, podemos escrever:

$$P(X=3)=\binom{5}{3}p^3q^{5-3}=\binom{5}{3}p^3q^2$$

Se a probabilidade de obtermos “cara” numa só prova (sucesso) é p = 1/2 e a probabilidade de não obtermos “cara” numa só prova (insucesso) é q = 1 – 1/2 = 1/2, então:

$$P(X=3)=\binom{5}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^2=\frac{5!}{3!\,2!}\left(\frac{1}{8}\right)\left(\frac{1}{4}\right)=\frac{5x4x3x2x1}{3x2x1x2x1}\left(\frac{1}{8}\right)\left(\frac{1}{4}\right)=\frac{5}{16}$$

Logo:

$$P(X=3)=\frac{5}{16}$$

2. Dois times de futebol, A e B, jogam entre si 6 vezes. Encontre a probabilidade do time A ganhar 4 jogos.

Solução:

Temos:

$N=6, k=4, p\ =\frac{1}{3}, q\ =1-\frac{1}{3}=\frac{2}{3}$

Então:

$$P(X=4)=\binom{6}{4}\left(\frac{1}{3}\right)^4\left(\frac{2}{3}\right)^2=15\left(\frac{1}{81}\right)\left(\frac{4}{9}\right)=\frac{20}{243}$$

Logo:

$$P(X=4)=\frac{20}{243}$$

EXERCÍCIOS

1. Determine a probabilidade de obtermos exatamente 3 caras em 6 lances de uma moeda.

2. Jogando-se um dado três vezes, determine a probabilidade de se obter um múltiplo de 3 duas vezes.

3. Dois times de futebol, A e B, jogam entre si 6 vezes. Encontre a probabilidade do time A:

a. ganhar dois ou três jogos;

b. ganhar pelo menos um jogo.

4. A probabilidade de um atirador acertar o alvo é 2/3. Se ele atirar 5 vezes, qual a probabilidade de acertar exatamente 2 tiros?

5. Seis parafusos são escolhidos ao acaso da produção de certa máquina, que apresenta 10% de peças defeituosas. Qual a probabilidade de serem defeituosos dois deles?

RESPOSTAS:

1. 5/32

2. 2/9

3. a. 400/729   b. 665/729

4. 40/243

5. 9,8415%

2. ENTENDENDO A FÓRMULA

O gerente da loja estima que de 10 vendas realizadas, 3 são microcomputadores e 7 equipamentos eletrônicos. Qual a probabilidade de que uma das próximas 4 vendas seja <u>um</u> microcomputador?

Começamos por determinar as 4 próximas vendas e depois suas probabilidades de ocorrência.

Sendo E a venda de um equipamento eletrônico e M a de um microcomputador, os quatro resultados possíveis (eventos elementares) são: EEEM, EEME, EMEE e MEEE.

Dos dados do gerente deduzimos que 70% das vendas realizadas são de equipamentos eletrônicos E e 30% de microcomputadores M. Se a sequência de venda de um M for EEEM sua probabilidade será igual a:

P(EEEM) = 0,70 x 0,70 x 0,70 x 0,30 = 0,30 x $0{,}70^3$

Aqui aplicamos a regra do produto, pois os eventos são independentes.

Aplicando o mesmo procedimento para os outros três eventos obteremos os mesmos resultados:

P(EEME) = 0,70 x 0,70 x 0,30 x 0,70 = 0,30 x $0{,}70^3$

P(EMEE) = 0,70 x 0,30 x 0,70 x 0,70 = 0,30 x $0{,}70^3$

P(MEEE) = 0,30 x 0,70 x 0,70 x 0,70 = 0,30 x $0{,}70^3$

Finalmente, como os quatro eventos são mutuamente excludentes, a probabilidade de que uma das quatro próximas vendas seja UM microcomputador é obtida pela regra da soma, assim:

P(x=1)= P(EEEM)+ P(EEME)+ P(EMEE) + P(MEEE)

Onde x = 1 identifica a venda de um microcomputador.

P(x=1) = 4 x (0,30 x $0{,}70^3$)= 0,4116 ou

P(x=1) = $\binom{4}{1}$ x $0{,}30^1$ x $0{,}70^3$ = 0,4116.

3. DISTRIBUIÇÃO BINOMIAL NO EXCEL

Vamos por meio de um exemplo fazer um histograma da distribuição binomial.

EXEMPLO 1. Uma experiência com distribuição binomial foi repetida 4 vezes seguidas. Considerando a probabilidade de sucesso p =0,50:

a. Calcule as probabilidades de todos os possíveis sucessos x.

b. Construa o gráfico da distribuição de probabilidades.

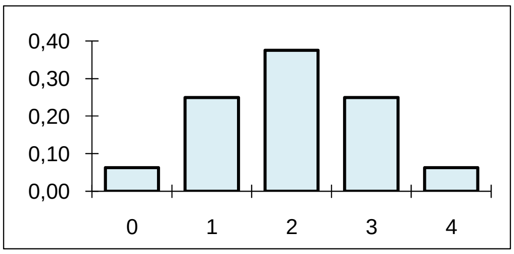
Solução:

Com a fórmula $P(X=k)=\binom{n}{k}p^k q^{n-k}$ construa uma planilha como a mostrada abaixo

	A	B	C	D	E	F	G	H	I	J	K
1	Exemplo										
2											
3		p	0,5								
4		n	4								
5		x	*p(x)*								
6		0	0,0625	<--=FATORIAL(C4)/(FATORIAL(B6)*FATORIAL(C4-B6))*C3^B6*(1-C3)^(C4-B6)							
7		1	0,2500	<--=FATORIAL(C4)/(FATORIAL(B7)*FATORIAL(C4-B7))*C3^B7*(1-C3)^(C4-B7)							
8		2	0,3750	<--=FATORIAL(C4)/(FATORIAL(B8)*FATORIAL(C4-B8))*C3^B8*(1-C3)^(C4-B8)							
9		3	0,2500	<--=FATORIAL(C4)/(FATORIAL(B9)*FATORIAL(C4-B9))*C3^B9*(1-C3)^(C4-B9)							
10		4	0,0625	<--=FATORIAL(C4)/(FATORIAL(B10)*FATORIAL(C4-B10))*C3^B10*(1-C3)^(C4-B10)							

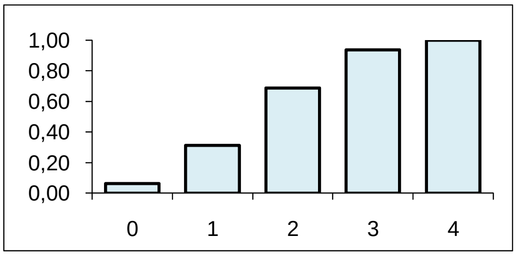
e a seguir com os dados da tabela construa o histograma:

	A	B	C	D	E	F	G
1	Exemplo						
2							
3		p	0,5				
4		n	4				
5		x	*p(x)*				
6		0	0,0625				
7		1	0,2500				
8		2	0,3750				
9		3	0,2500				
10		4	0,0625				
11							



Continuando, podemos calcular a probabilidade de que x seja menor que 2 e de que x seja menor ou igual a 2. Para isto construímos a tabela e o gráfico de probabilidades acumuladas mostrados abaixo, onde temos que P(x<2) = 0,3125 e P(x ≤ 2) = 0,6875

π	0,5
n	4
x	*p(x)*
0	0,0625
1	0,3125
2	0,6875
3	0,9375
4	1,0000



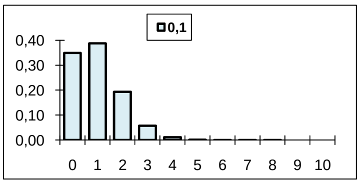
EXEMPLO 2

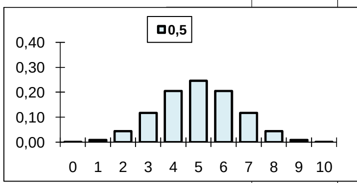
Uma experiência com distribuição binomial foi repetida 10 vezes seguidas. Construa a tabela completa de probabilidades e o histograma de x considerando quatro valores de probabilidades de sucesso p = 0,10, p = 0,50, p = 0,70 e p = 1.

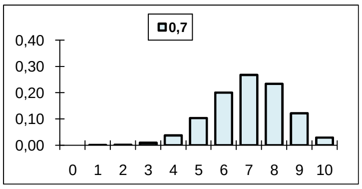
Solução:

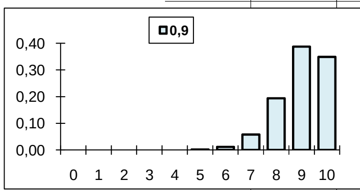
Exemplo

n	10			
p	0,1	0,5	0,7	0,9
x	*p(x)*	*p(x)*	*p(x)*	*p(x)*
0	0,3487	0,0010	0,0000	0,0000
1	0,3874	0,0098	0,0001	0,0000
2	0,1937	0,0439	0,0014	0,0000
3	0,0574	0,1172	0,0090	0,0000
4	0,0112	0,2051	0,0368	0,0001
5	0,0015	0,2461	0,1029	0,0015
6	0,0001	0,2051	0,2001	0,0112
7	0,0000	0,1172	0,2668	0,0574
8	0,0000	0,0439	0,2335	0,1937
9	0,0000	0,0098	0,1211	0,3874
10	0,0000	0,0010	0,0282	0,3487









A tabela abaixo fornece a probabilidade de ocorrerem x sucessos em n experiências com probabilidades de sucesso definidas na própria tabela.

TABELA DA DISTRIBUIÇÃO BINOMIAL

n 7 Probabilidade de X

x	0,05	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	0,95
0	0,6983	0,4783	0,2097	0,0824	0,0280	0,0078	0,0016	0,0002	0,0000	0,0000	0,0000
1	0,2573	0,3720	0,3670	0,2471	0,1306	0,0547	0,0172	0,0036	0,0004	0,0000	0,0000
2	0,0406	0,1240	0,2753	0,3177	0,2613	0,1641	0,0774	0,0250	0,0043	0,0002	0,0000
3	0,0036	0,0230	0,1147	0,2269	0,2903	0,2734	0,1935	0,0972	0,0287	0,0026	0,0002
4	0,0002	0,0026	0,0287	0,0972	0,1935	0,2734	0,2903	0,2269	0,1147	0,0230	0,0036
5	0,0000	0,0002	0,0043	0,0250	0,0774	0,1641	0,2613	0,3177	0,2753	0,1240	0,0406
6	0,0000	0,0000	0,0004	0,0036	0,0172	0,0547	0,1306	0,2471	0,3670	0,3720	0,2573
7	0,0000	0,0000	0,0000	0,0002	0,0016	0,0078	0,0280	0,0824	0,2097	0,4783	0,6983

A tabela abaixo mostra a probabilidade acumulada de ocorrerem até x sucesso em n experiências com as probabilidades de sucesso definidas na própria tabela.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	**TABELA DA DISTRIBUIÇÃO BINOMIAL**												
2													
3		***n***	**7**	Probabilidade Acumulada									
4													
5		***x***	**0,05**	**0,1**	**0,2**	**0,3**	**0,4**	**0,5**	**0,6**	**0,7**	**0,8**	**0,9**	**0,95**
6		**0**	**0,6983**	**0,4783**	**0,2097**	**0,0824**	**0,0280**	**0,0078**	**0,0016**	**0,0002**	**0,0000**	**0,0000**	**0,0000**
7		**1**	**0,9556**	**0,8503**	**0,5767**	**0,3294**	**0,1586**	**0,0625**	**0,0188**	**0,0038**	**0,0004**	**0,0000**	**0,0000**
8		**2**	**0,9962**	**0,9743**	**0,8520**	**0,6471**	**0,4199**	**0,2266**	**0,0963**	**0,0288**	**0,0047**	**0,0002**	**0,0000**
9		**3**	**0,9998**	**0,9973**	**0,9667**	**0,8740**	**0,7102**	**0,5000**	**0,2898**	**0,1260**	**0,0333**	**0,0027**	**0,0002**
10		**4**	**1,0000**	**0,9998**	**0,9953**	**0,9712**	**0,9037**	**0,7734**	**0,5801**	**0,3529**	**0,1480**	**0,0257**	**0,0038**
11		**5**	**1,0000**	**1,0000**	**0,9996**	**0,9962**	**0,9812**	**0,9375**	**0,8414**	**0,6706**	**0,4233**	**0,1497**	**0,0444**
12		**6**	**1,0000**	**1,0000**	**1,0000**	**0,9998**	**0,9984**	**0,9922**	**0,9720**	**0,9176**	**0,7903**	**0,5217**	**0,3017**
13		**7**	**1,0000**	**1,0000**	**1,0000**	**1,0000**	**1,0000**	**1,0000**	**1,0000**	**1,0000**	**1,0000**	**1,0000**	**1,0000**
14													

4. ESPERANÇA, VARIÂNCIA e DESVIO PADRÃO do MODELO BINOMIAL

Aplicando os conceitos de valor esperado nas distribuições discretas, substituindo a expressão P(x) da distribuição binomial naquelas expressões obteremos o *valor esperado* E(x) = μ, a variância Var(X) = σ^2 e o desvio padrão σ da distribuição binomial. Perceba o leitor que estes resultados não dependem do número de sucessos x.

Parâmetros da distribuição binomial

A média, a variância e o desvio padrão são obtidos com:

$\mu = n \times p$ $\sigma^2 = n \times p \times (1-p)$ e $\sigma = \sqrt{n\, x\, p\, x\, (1-p)}$

EXEMPLO 3

São realizadas 10 experiências com probabilidade de sucesso p = 0,10. Considerando que o experimento tem distribuição binomial, calcular a média e o desvio padrão

Solução:

Aplicando as fórmulas temos:

$\mu = n \times p = 10 \times 0{,}1 = 1$

$\sigma = \sqrt{n\, x\, p\, x\, (1-p)} = \sqrt{10\, x\, 0{,}10\, x\, (1-0{,}10)} = 0{,}9487$

EXEMPLO 4

Você tem uma carteira com 15 ações. No pregão de ontem 75% das ações na bolsa de valores caíram de preço. Supondo que as ações que perderam valor têm distribuição binomial:

- Quantas ações da sua carteira você espera que tenham caído de preço?
- Qual o desvio padrão das ações que tem na carteira?
- Qual a probabilidade que as 15 ações da carteira tenham caído?
- Qual a probabilidade que tenham caído de preço exatamente 10 ações?
- Qual a probabilidade que treze ou mais ações tenham caído de preço?

Solução:

Como 75% das ações caíram de preço, o número de ações da carteira que devem ter caído de preço será 11,25 = 0,75 x 15. O desvio padrão foi:

$\sigma = \sqrt{n\, x\, p\, x\, (1-p)} = \sqrt{15\, x\, 0{,}75\, x\, (1-0{,}75)} = 1{,}67$

$$P(X=15) = \binom{n}{k} p^k q^{n-k} = \binom{15}{15} 0{,}75^{15}(1-0{,}75)^{15-15} = \frac{15!}{13!\,(15-15)!}(0{,}75)^{15}(0{,}25)^0 = 0{,}0134$$

De forma equivalente, a probabilidade que tenham caído de preço exatamente 10 ações é P(x = 10) = 0,1651, e a probabilidade que treze ou mais ações tenham caído de preço é obtida com P(x ≥ 13) = P(x = 13) + P(x = 14) + P(x = 15) = 0,2361

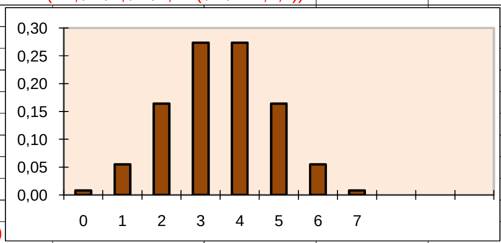
O Excel dispõe de funções estatísticas para realizar cálculos com a distribuição normal. As sintaxes dessas funções são as seguintes:

DISTRBINOM(num_s;tentativas;probabilidade_s;cumulativo)

Esta função dá a probabilidade ou a probabilidade acumulada do num_s conforme o valor do argumento cumulativo.

- Se o argumento cumulativo for FALSO, a função dará a probabilidade do número de sucessos num_s com probabilidade_s de sucesso para um número de tentativas independentes.
- Se o argumento cumulativo for VERDADEIRO, a função dará a probabilidade acumulada do número máximo de sucessos num_s com probabilidade_s de sucesso para um número de tentativas independentes.

	A	B	C	D	K
1	PROBABILIDADES DA DISTRIBUIÇÃO BINOMIAL				
2					
3		p	0,5		Probabilidade de X
4		n	7		Probabilidade Acumulada
5		μ	3,5	<--=C4*C3	1
6		σ	1,32	<--=RAIZ(C4*C3*(1-C3))	
7		Probabilidade de X			
8		x	P(=x)		
9		0	0,0078	<--=DISTRBINOM(B9;C4;C3;SE(K5=1;0;1))	
10		1	0,0547		
11		2	0,1641		
12		3	0,2734		
13		4	0,2734		
14		5	0,1641		
15		6	0,0547		
16		7	0,0078		
17					
18					
19					
20		<--=SE(OU(B10=C4;B10="");"";1+B10)			
21					



Comparando a parte teórica com a função DISTRBINOM teremos:

- Se em n experiências com distribuição binomial acontecerem x sucessos com probabilidade p, a probabilidade de ocorrerem x sucessos P(x) será obtida com a função estatística:

  DISTRBINOM(num_s;tentativas;probabilidade_s;FALSO)

  Esta função corresponde à expressão: $P(X) = \binom{n}{x} p^x q^{n-x}$, para x = 0, 1, 2, ..., n.

- Se em n experiências com distribuição binomial acontecerem x sucessos com probabilidade p, a probabilidade acumulada de ocorrerem até x sucessos P(x) será obtida com a função estatística:

  DISTRBINOM(num_s;tentativas;probabilidade_s;VERDADEIRO)

  Esta função corresponde à expressão:

$$P(x) = \sum_{i=0}^{x} \binom{n}{i} p^i (1-p)^{n-i}$$

Na Figura acima, selecionando a opção de cálculo na caixa de combinação do modelo, você poderá calcular a probabilidade de x sucessos e a probabilidade acumulada até x sucessos de n = 10 repetições do experimento.

EXEMPLO 5

Seja uma experiência com distribuição binomial com n = 4 e a probabilidade de sucesso p = 0,3. Calcular a probabilidade de ter 2 sucesso e a probabilidade de ter 2 sucessos.

Solução:

A probabilidade de ter 2 sucessos é P(x=1) = 0,2646, valor obtido com a fórmula:

= DISTRBINOM(2;4;0,3;FALSO)

Da mesma maneira a probabilidade de ter até 2 sucessos é P(x ≤ 2) = 0,9163, valor obtido com a fórmula:

= DISTRBINOM(2;4;0,3;VERDADEIRO).

Aqui vai um segmento de planilha que realiza este cálculo:

	A	B	C	D
22	**CÁLCULO DE PROBABILIDADES BINOMIAIS**			
23				
24		**p**	0,3	
25		***n***	4	
26		***x***	2	
27		***P(x=2)***	0,2646	
28		***P(x<=2)***	0,9163	
29		**µ**	1,2	
30		**σ**	0,92	
31				

3.1 - OUTRAS FUNÇÕES ESTATÍSTICAS ASSOCIADAS À DISTRIBUIÇÃO BINOMIAL

PROB(intervalo_x;intervalo_prob;limite_inferior;limite_superior)

A função estatística PROB dá a **probabilidade acumulada** entre o *limite inferior* e o *limite superior*, ambos incluídos, do intervalo_x de valores e o intervalo_prob de probabilidades associadas aos valores x.

A figura abaixo mostra um modelo em que utilizamos a função PROB com os dados do Exemplo 5

	B	C	D	E	F	G	H	I	J	K	L
33	**ıção PROB**										
34											
35	**p**	0,3									
36	***n***	4									
37	***x***	***P(=x)***									
38	0	0,2401	**Limite inferior**	1							
39	1	0,4116	**Limite superior**	3							
40	2	0,2646	**Prob. Acumulada**	0,7518	<--=PROB(B38:B42;C38:C42;E38;E39)						
41	3	0,0756	**PROB - matriz**	0,7518	<--=PROB({0;1;2;3;4};{0,2401;0,4116;0,2646;0,0756;0,0081};E38;E39)						
42	4	0,0081	**ɔom DISTRBINOM**	0,7518	<--=DISTRBINOM(E39;C36;C35;VERDADEIRO)-SE(E38=0;0;DISTRBINOM(E38-1;C36;C35;VERDADEIRO))						
43											

- No intervalo B38:B42 foram registrados os valores de x, e no intervalo C38:C42 foram calculadas as probabilidades correspondentes, como mostra a figura acima.
- No intervalo E38:E39 foram registrados o limite inferior e o limite superior de x, respectivamente, valores 1 e 3.
- Na célula E40, com a fórmula: =PROB(B38:B42;C38:C42;E38:E39) foi calculada a probabilidade acumulada $P(1 \leq x \leq 3) = 0,8448$. Verifique que a probabilidade acumulada $P(1 \leq x \leq 3) = P(x \leq 3) - P(x = 0) = 0,8704 - 0,0256 = 0,8448$.
- O mesmo resultado é obtido informando os dados em forma de matriz, registrando na célula E41 a fórmula:

  =PROB({0;1;2;3;4};{0,2401;0,4116;0,2646;0,0756;0,0081};E38;E39)
- Com a função DISTRBINOM, registrando na célula E42 a fórmula:

  =DISTRBINOM(E39;C36;C35;VERDADEIRO)-SE(E38=0;0;DISTRBINOM(E38-1;C36;C35;VERDADEIRO))

  Perceba que ao valor do num_s da segunda parcela da fórmula foi subtraído um. Entretanto, quando o limite inferior de x for zero, o argumento num_s da segunda parcela da fórmula acima será zero.

CRIT.BINOM(tentativas;probabilidade_s;alfa)

A função estatística CRIT.BINOM dá o **menor número de sucessos** para o qual a distribuição binomial *acumulada* é <u>maior ou igual</u> ao argumento alfa. Por exemplo, com os dados do Exemplo 5 , se alfa = 0,50 o número de sucessos menor ou igual a 0,50 é dois, como mostra a figura abaixo.

	A	B	C	D	E	F
45	**Função CRIT.BINOM**					
46						
47		**p**	0,3			
48		***n***	4			
49		***x***	***P(<=x)***			
50		0	0,2401	<--=DISTRBINOM(B50;C48;C47;VERDADEIRO		
51		1	0,6517			
52		2	0,9163			
53		3	0,9919			
54		4	1,0000			
55		***alfa***	0,6000			
56		***x***	1	<--=CRIT.BINOM(C48;C47;C55)		
57						

Por exemplo, a função CRIT.BINOM determina o número máximo de peças defeituosas de um lote de produção sem rejeitar o lote inteiro. Para valores exatos de probabilidade acumulada, a função estatística CRIT.BINOM é inversa da função estatística DISTRBINOM com o argumento cumulativo VERDADEIRO.

Distribuição de Poisson

A *distribuição de Poisson* é empregada em experimentos, nos quais não se está interessado no número de sucessos obtidos em *n* tentativas, como ocorre no caso da distribuição Binomial, mas sim no número de sucessos ocorridos durante um **intervalo contínuo**, que pode ser um intervalo de tempo, espaço, etc. Como por exemplo:

- O número de suicídios ocorridos em uma cidade durante um ano;
- O número de acidentes automobilísticos ocorridos numa rodovia em um mês;
- Número de chegadas a um caixa automático de um banco durante um período de 15 minutos
- A probabilidade de um carro chegar a um posto de gasolina em quaisquer dois períodos de tempo de mesmo tamanho.
- A chegada ou não chegada de um carro em qualquer período de tempo independentemente da chegada ou não chegada de outro carro em qualquer outro período.
- Defeitos por unidade (m^2, m, etc.) por peça fabricada
- Erros tipográficos por página, em um material impresso
- Carros que passam por um cruzamento por minuto, durante certa hora do dia.
- Usuários de computador ligados à Internet

Note que nos exemplos acima, não há como determinar-se a probabilidade de ocorrência de um sucesso, mas sim a frequência média de sua ocorrência, como, por exemplo, dois suicídios por ano, a qual será que denominada λ.

É, então, uma distribuição de probabilidade **discreta** que se aplica a ocorrência de eventos ao longo de <u>intervalos</u> especificados. A *variável aleatória* é o <u>número de ocorrência do evento no intervalo</u>. Os intervalos podem ser de tempo, distância, área, volume ou alguma unidade similar.

Uma variável aleatória **X** admite *distribuição de Poisson* se:

1. X = {0, 1, 2, ... }(não tem limites);

2. $P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$, k = 0, 1, 2, ...; é a probabilidade de **k** ocorrências em um intervalo

3. $E(X) = \mu = \lambda$;

4. $Var(X) = \sigma^2 = \lambda$.

Prova das propriedades 3 e 4:

$$E(X) = \sum_{x=0}^{n} xP(X = x) = \sum_{x=0}^{n} x\frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=0}^{n} \frac{e^{-\lambda}\lambda^x}{(x-1)!} = \sum_{s=-1}^{n-1} \frac{e^{-\lambda}\lambda^{s+1}}{s!} = \lambda\sum_{s=-1}^{n-1} \frac{e^{-\lambda}\lambda^s}{s!} = \lambda$$

$$E(X^2) = \sum_{x=0}^{n} x^2P(X = x) = \sum_{x=0}^{n} x^2\frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=0}^{n} x\frac{e^{-\lambda}\lambda^x}{(x-1)!} = \sum_{s=-1}^{n-1} (s+1)\frac{e^{-\lambda}\lambda^{s+1}}{s!} = \lambda\sum_{s=-1}^{n-1} (s+1)\frac{e^{-\lambda}\lambda^s}{s!}$$
$$= \lambda\left[\sum_{s=-1}^{n-1} s\frac{e^{-\lambda}\lambda^s}{s!} + \sum_{s=-1}^{n-1} \frac{e^{-\lambda}\lambda^s}{s!}\right] = \lambda[\lambda + 1] = \lambda^2 + \lambda$$

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

Uma *distribuição de Poisson* difere de uma *distribuição binomial* nestes aspectos fundamentais:

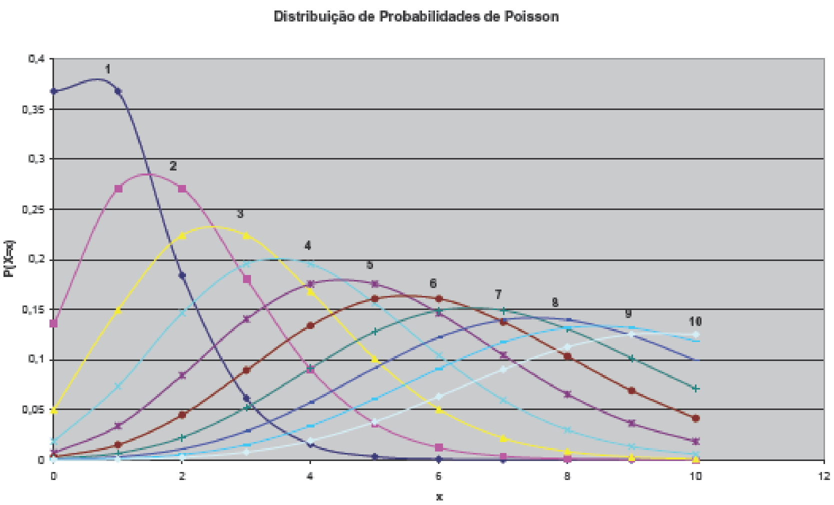
1. A distribuição binomial é afetada pelo tamanho da amostra n e pela probabilidade p, enquanto que a distribuição de Poisson é afetada apenas pela média _;

2. Na distribuição binomial, os valores possíveis da variável aleatória X são 0; 1; 2; _ _ _; n, mas a distribuição de Poisson têm os valores de X de 0; 1; 2; _ _ _, sem qualquer limite superior.

Obs: O parâmetro λ é usualmente referido como <u>taxa de ocorrência</u>.

Propriedades do experimento de Poisson:

- A probabilidade de uma ocorrência é a mesma para quaisquer dois intervalos
- A ocorrência ou não ocorrência em qualquer intervalo é independente da ocorrência ou não-ocorrência em qualquer intervalo.



EXEMPLO 1

O Corpo de Bombeiros de uma determinada cidade recebe, em média, 3 chamadas por dia. Qual a probabilidade de receber:

a) 4 chamadas num dia

λ = 3 chamadas por dia em média

$P(X = 4) = \frac{e^{-\lambda}\lambda^k}{k!} = \frac{e^{-3}3^4}{4!} = 0{,}1680 \; ou \; 16{,}80\%$

b) Nenhuma chamada em um dia

$P(X = 4) = \frac{e^{-\lambda}\lambda^k}{k!} = \frac{e^{-3}3^0}{0!} = 0{,}0498 \; ou \; 4{,}98\%$

c) 20 chamadas em uma semana.

X = número de chamadas por dia

Y = número de chamadas por semana

$E(X) = \lambda = 3$ chamadas por dia $\Rightarrow E(Y) = \lambda^* = 7 \times E(X) = 21$ chamadas por semana.

$P(Y = 20) = \frac{e^{-\lambda}\lambda^k}{k!} = \frac{e^{-21}21^{20}}{20!} = 0{,}0867 \; ou \; 8{,}67\%$

EXEMPLO 2

Uma central telefônica tipo PABX recebe uma *média* de 5 chamadas por minuto. Qual a probabilidade deste PABX não receber nenhuma chamada durante um intervalo de 1 minuto?

$$P(X=0) = \frac{5^0 . e^{-5}}{0!} = e^{-5} = 0{,}0067$$

X = v. a. n° de chamadas em um intervalo de tempo
λ= taxa de ocorrência de chamadas (n° esperado de chamadas)

Aproximação da distribuição Binomial a Poisson.

Pode-se demonstrar que uma distribuição Binomial, cujo evento de interesse (sucesso) é raro (p muito pequeno e n muito grande), tende para uma distribuição de Poisson. Na prática, a aproximação é considerada boa quando $n \geq 50$ e $p \leq 0{,}10$.

Aproximação: Sabe-se que se $X \sim B(n; p)$, $E(X) = np$, então $, = E(X) = np$

EXEMPLO 3

A probabilidade de um indivíduo sofrer uma reação alérgica, resultante da injeção de determinado soro é de 0,01. Determinar a probabilidade de entre 200 indivíduos, submetidos a este soro, nenhum sofrer esta reação alérgica.

$X \sim B(200; 0,01) \Rightarrow E(X) = n.p = 200x0,01 = 2 = \lambda$

$P(X = 2) \cong \frac{e^{-\lambda}\lambda^k}{k!} = \frac{e^{-2}2^0}{0!} = 0,1353\ ou\ 13,53\%$

2. DISTRIBUIÇÃO DE POISSON NO EXCEL

O Excel dispõe da função estatística POISSON cuja sintaxe é:

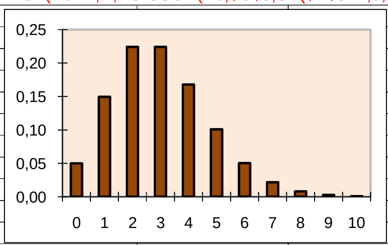
POISSON(x;média;cumulativo)

A função estatística POISSON dá a probabilidade ou a probabilidade acumulada conforme o valor do argumento *cumulativo*:

- Se o argumento cumulativo for FALSO a função dará a probabilidade de x considerando a média. O resultado P(x=4) = 16,80% é obtido com a fórmula: = POISSON(4;3;FALSO).
- Se o argumento cumulativo for VERDADEIRO a função dará a probabilidade acumulada até x considerando a média. O resultado P(x≤4)= 81,53% foi obtida com a fórmula: =POISSON(4;3;VERDADEIRO).

A Figura mostra o modelo **Probabilidades da Distribuição de Poisson** com os dados do Exemplo 1. Selecionando a opção de cálculo na caixa de combinação do modelo pode-se calcular a <u>probabilidade</u> de **x** ocorrências e a <u>probabilidade acumulada</u> de até **x** ocorrências

	A	B	C	D	E	F	G	H	I	J	K
1	PROBABILIDADES DA DISTRIBUIÇÃO DE POISSON										
2											
3		Média	3,00								Probabilidade de X
4		Probabilidade de X									Probabilidade Acumulada
5		x	P(=x)								1
6		0	0,0498	<--=SE(B6="";"";POISSON(B6;C3;SE(K5=1;0;1)))							
7		1	0,1494								
8		2	0,2240								
9		3	0,2240								
10		4	0,1680								
11		5	0,1008								
12		6	0,0504								
13		7	0,0216								
14		8	0,0081								
15		9	0,0027								
16		10	0,0008								
17											
18											



EXEMPLO 4

O erro de digitação cometido pelos caixas é 0,35 por hora. Qual a probabilidade de que um caixa cometa 2 erros numa hora?

Solução

A probabilidade P(x = 2)= 4,32 é obtida com a fórmula de distribuição de Poisson:

$P(X = 2) \cong \frac{e^{-\lambda}\lambda^k}{k!} = \frac{e^{-0,35}0,35^2}{2!} = 0,04316\ ou\ 4,316\%.$

A Figura abaixo mostra o cálculo realizado na planilha

	A	B	C	D	E	F
19	CÁLCULO DE PROBABILIDADES - POISSON					
20						
21		Média	0,35			
22		x	2			
23		P(x=2)	0,0432	<--=POISSON(C22;C21;FALSO)		
24		P(x<=2)	0,9945	<--=POISSON(C22;C21;VERDADEIRO)		
25						

Distribuição Geométrica

Suponha-se um experimento, no qual estamos interessados apenas na ocorrência ou não de um determinado evento, como, por exemplo, o sexo do filho de uma determinada mulher ser feminino. E, assim como na distribuição binomial, que esse experimento seja repetido um número n de vezes, que em cada repetição seja independente das demais e que a probabilidade de sucesso p em cada repetição seja constante. Suponha-se que o experimento seja repetido até que ocorra o primeiro sucesso (o sexo do filho seja feminino).

Então a variável aleatória: X = número de tentativas até que se obtenha o primeiro sucesso, seguirá uma *distribuição geométrica*, com parâmetro p (probabilidade de sucesso) . Simbolicamente $X \sim G(p)$.

Função de Probabilidade

Como o experimento será repetido até que se obtenha o primeiro sucesso, e considerando que esse ocorra na k-ésima repetição, deverão ocorrer k -1 fracassos antes que o experimento seja encerrado. Assim, a probabilidade de que a variável aleatória X = número de repetições até se obter o primeiro sucesso é:

$$P(X = x) = pq^{x-1}$$

com
p = probabilidade de “sucesso";
$q = 1 - p$ = probabilidade de “fracasso"

Parâmetros característicos

$$E(X) = \frac{1}{p}$$
$$Var(X) = \frac{q}{p^2}$$

EXEMPLO 1

Um casal com problemas para engravidar, recorreu a uma técnica de inseminação artificial no intuito de conseguir o primeiro filho. A eficiência da referida técnica é de 0,20 e o custo de cada inseminação U\$2000,00.

a) Qual a probabilidade de que o casal obtenha êxito na terceira tentativa?

$$P(X = k) = pq^{k-1} = (0{,}2)(0{,}8)^3 = 0{,}128 \; ou \; 12{,}80\%$$

b) Qual o custo esperado deste casal para obter o primeiro filho?

$$E(X) = \frac{1}{p} = \frac{1}{0{,}2} = 5$$

Custo esperado = 5 x 2000,00 = U\$10.000, 00

EXEMPLO 2

Bob é o jogador de basquete da faculdade. Ele é um lançador de arremessos livres 70%. Isto significa que sua probabilidade de acertar um arremesso livre é 0,70. Durante uma partida, qual é a probabilidade que Bob acerte seu primeiro arremesso livre no seu quinto arremesso?

Solução

Este é um exemplo de uma distribuição geométrica, que como veremos é um caso especial de uma distribuição binomial negativa. Logo, usando a fórmula da distribuição geométrica termos:

$$P(X = k) = pq^{k-1} = (0{,}7)(0{,}3)^4 = 0{,}00567 \; ou \; 0{,}567\%$$

Distribuição Binomial Negativa[1]

Nas mesmas condições em que foi definida a distribuição geométrica, e considerando que o experimento será repetido até que se obtenha o *r-ésimo* sucesso, então a variável *X = número de tentativas até se obter o r-ésimo sucesso* seguirá a distribuição binomial negativa.

Um experimento binomial negativo é um experimento estatístico que tem as seguintes propriedades:

- O experimento consiste de x tentativas repetidas.
- Cada tentativa pode resultar em apenas dois resultados possíveis. Podemos chamar um destes resultados de sucesso e o outro de fracasso.
- A probabilidade de sucesso, denotada por p, é a mesma em cada tentativa.
- As tentativas são independentes; isto é, o resultado de uma tentativa não afeta o resultado das outras tentativas.
- O experimento continua até que r sucessos sejam observados, onde r é especificado antecipadamente.

Considere o seguinte experimento estatístico. Você lança uma moeda repetidamente e conta o número de vezes que sai cara como resultado. Você continua lançando a moeda até que tenha saído 5 vezes cara. Este é um experimento binomial negativo porque:

- O experimento consiste de tentativas repetidas. Lançamos uma moeda repetidamente até que cara tenha saído 5 vezes.
- Cada tentativa pode resultar em apenas dois resultados possíveis – cara ou coroa.
- A probabilidade de sucesso é constante – 0,5 em cada tentativa.
- As tentativas são independentes; isto é, obter cara numa tentativa não afeta se obteremos cara nas outras tentativas.
- O experimento continua até que um número fixo de sucessos tenha ocorrido; neste caso,5 caras.

NOTAÇÃO

A seguinte notação é útil, quando falamos a respeito da probabilidade binomial negativa:

- K: O número de tentativas exigido para se produzir r sucessos num experimento binomial negativo.
- r: O número de sucessos no experimento binomial negativo.
- p: A probabilidade de sucesso numa tentativa individual.
- q: A probabilidade de fracasso numa tentativa individual. (Isto é igual a 1 – p).
- $b^*(k;r,p)$: Probabilidade binomial negativa - a probabilidade que um experimento binomial negativo de x-tentativas resulte em r sucessos na k-ésima tentativa, quando a probabilidade de sucesso na tentativa individual é p.
- $C_{(r)}^{(n)}$: O número de combinações de n coisas, tomando r coisas de cada vez.

Variável aleatória binomial negativa

Uma **variável aleatória binomial negativa** é o número X de tentativas repetidas para produzir **r** sucessos num *experimento* binomial negativo. A distribuição de probabilidade de uma variável aleatória binomial negativa é chamada de **distribuição binomial negativa.**

Suponha que lancemos uma moeda repetidamente e contemos o número de caras (sucessos). Se continuarmos lançando a moeda até que tenha saído cara 2 vezes, estamos conduzindo um experimento binomial negativo. A variável aleatória binomial negativa é o número de lançamentos exigidos para se conseguir cara 2 vezes. Neste exemplo, o número de moedas lançadas é uma variável aleatória que pode assumir qualquer valor inteiro entre 2 e $+\infty$. A distribuição de probabilidade binomial negativa para este exemplo é apresentada abaixo:

Número de Moedas Lançadas	Probabilidade
2	0,25
3	0,25
4	0,1875
5	0,125
6	0,078125
7 ou mais	0,109375

Função de Probabilidade

[1] Também conhecida como distribuição de Pascal

Para que o *r-ésimo* sucesso ocorra na *k-ésima* tentativa é necessário que <u>ocorra</u> um sucesso nesta tentativa (repetição do experimento) e que tenham ocorridos $(r-1)$ sucessos nas $(k-1)$ repetições anteriores[2]. Dado que a probabilidade de ocorrência de sucesso, numa dada repetição do experimento é dada por p e a probabilidade de ocorrerem $r-1$ sucessos em k - 1 repetições, sendo estes dois eventos independentes, a probabilidade de que o *r-ésimo* sucesso ocorra na *k-ésima* repetição do experimento é dada por:

$$b^{*}(X=k;r,p) = p\,.[C_{(r-1)}^{(k-1)}p^{r-1}q^{(k-1)-(r-1)}] = C_{(r-1)}^{(k-1)}p^{r}q^{k-r} \quad ; k \geq r$$

onde:

p = probabilidade de "sucesso"; q = 1 - p = probabilidade de "fracasso"

Parâmetros característicos:

$$E(X) = \frac{r}{p}$$

$$Var\,(X) = \frac{rq}{p^2}$$

<u>EXEMPLO 1</u>

Bob é o jogador de basquete da faculdade. Ele é um lançador de arremessos livres 70%. Isto significa que sua probabilidade de acertar um arremesso livre é 0,70. Durante uma partida, qual é a probabilidade que Bob acerte seu terceiro arremesso livre no seu quinto arremesso?

Solução

Este é um exemplo de um experimento binomial negativo. A probabilidade de sucesso (p) é 0,70, o número de tentativas (k) é 5, e o número de sucessos r é 3.
Para resolver este problema, entremos com estes valores na fórmula (fmp)da binomial negativa

$$b^{*}(X=5;3,0{,}7) = C_{(2)}^{(4)}0{,}7^{3}0{,}3^{5} = 6\,.0{,}343\,.0{,}09 = 0{,}18522$$

2. DISTRIBUIÇÃO BINOMIAL NEGATIVA NO EXCEL

O Excel dispõe da função estatística DIST.BIN.NEG cuja sintaxe é:

DIST.BIN.NEG(num_f;num_s;probabilidade_s)

Esta função dá a probabilidade de acontecer o número determinado de falhas ou insucesso (num_f = k-r) antes de acontecer um número **r** de sucessos (num_s) com probabilidade de sucesso (probabilidade_s) constante.

Por exemplo, a probabilidade de ocorrerem 4 falhas antes de acontecerem 3 sucessos com probabilidade de sucesso constante 0,40 é igual a 12,44%, valor obtido com a fórmula: = DIST.BIN.NEG(4;3;0,4) na planilha abaixo:

	A	B	C	D	E
1	**Função DIST.BIN.NEG**				
2					
3		p	0,4		
4		*x*	3		
5		*não x*	*P (x)*		
6		0	0,0640		
7		1	0,1152		
8		2	0,1382		
9		3	0,1382		
10		4	0,1244	<--=DIST.BIN.NEG(B10;C4;C3)	

É fácil de verificar que se o número de falhas for 0, a função DIST.BIN.NEG dá o mesmo resultado da função BINOMDIST, considerando que o número de experimentos seja igual ao número de sucessos e o argumento cumulativo FALSO: DIST.BIN.NEG(0;2;0,40 = DISTRBINOM(2;2;0,40;Falso).

[2] No exemplo anterior, vemos, pela tabela, que a probabilidade binomial negativa de se obter a segunda cara no sexto lançamento da moeda é 0,078125.

Distribuição Hipergeométrica

Um experimento hipergeométrico é um experimento estatístico que tem as seguintes propriedades:

- Uma amostra de tamanho n é selecionada aleatoriamente sem reposição de uma população de N itens.
- Na população, k itens podem ser classificados como sucessos e N – k itens podem ser classificados como fracassos.

Considere o seguinte experimento estatístico. Você tem uma urna de 10 bolinhas de gude – 5 vermelhas e 5 verdes. Você seleciona aleatoriamente 2 bolinhas de gude sem reposição e conta o número de bolinhas vermelhas que você selecionou. Este seria um experimento hipergeométrico.

Note que não será um experimento binomial. Um experimento binomial exige que a probabilidade de sucesso seja constante em cada tentativa. Com o experimento acima, a probabilidade de um sucesso muda em cada tentativa. No início, a probabilidade de selecionar uma bolinha vermelha é 5/10. Se você selecionar uma bolinha vermelha na primeira tentativa, a probabilidade de selecionar uma bolinha vermelha na segunda tentativa é 4/9. E se você selecionar uma bolinha verde na primeira tentativa, a probabilidade de selecionar uma bolinha vermelha na segunda tentativa é 5/9.
Note ainda que se você selecionou as bolinhas com reposição, a probabilidade de sucesso não mudaria. Ela seria 5/10 em cada tentativa. Então, este seria um experimento binomial.

NOTAÇÃO

A seguinte notação é útil, quando falamos a respeito da probabilidade hipergeométrica e distribuições hipergeométricas:

- N: O número de itens na população.
- k: O número de itens na população que são classificados como sucessos.
- n: O número de itens na amostra.
- X: O número de itens na amostra que são classificados como sucessos.
- $C_{(x)}^{(k)}$: O número de combinações de k coisas, tomando x coisas de cada vez.
- h(x;N,n,k): Probabilidade hipergeométrica - a probabilidade que um experimento hipergeométrico de n-tentativas resulte em exatamente x sucessos, quando população consistir de N itens, k dos quais são classificados como sucessos.

Função de Probabilidade

Uma variável aleatória hipergeométrica X é o número de sucessos que resulta de um experimento hipergeométrico. A distribuição de probabilidades de uma variável aleatória hipergeométrica é chamada função distribuição hipergeométrica.

$$h(X=x;N,n,k)=\frac{C_x^k C_{(n-x)}^{(N-k)}}{C_n^N}$$

Parâmetros característicos:
Fazendo $\frac{k}{N}=p \quad e \quad \frac{N-k}{N}=q$ tem-se

$$\mathrm{E(X)} = \mathrm{n} \,.\, \mathrm{p}$$
$$\mathrm{Var(X)} = n.p.q.\frac{N-n}{N-1}$$

EXEMPLO 1

No fichário de um hospital, estão arquivados os prontuários dos de 20 pacientes, que deram entrada no PS apresentando algum problema cardíaco. Destes 5 sofreram infarto.
Retirando-se uma amostra ao acaso de 3 destes prontuários, qual a probabilidade de que dois deles sejam de pacientes que sofreram infarto?

Solução:

N = 20 k = 5 n = 3 X = 2

$$h(X=2;20{,}3{,}5)=\frac{C_2^5 C_{(3-2)}^{(20-5)}}{C_3^{20}}=\frac{C_2^5 C_{(1)}^{(15)}}{C_3^{20}}=\frac{10\ x\ 15}{1140}=0{,}1315\ ou\ 13{,}15\%$$

EXEMPLO 2

Suponha que selecionemos aleatoriamente 5 cartas baralho sem reposição de um de um maço ordinário de jogo de baralho. Qual é a probabilidade de obter exatamente 2 cartas de baralho vermelhas (isto é, copas ou ouros)?

Solução:

N = 52   k = 26 cartas vermelhas   n = 5 cartas selecionadas aleatoriamente X = 2

$$h(X=2;52,5,26)=\frac{C_2^{26}C_{(5-2)}^{(52-26)}}{C_5^{52}}=\frac{C_2^{26}C_{(3)}^{(26)}}{C_5^{52}}=\frac{325\ x\ 2.600}{2.598.960}=0{,}32513\ ou\ 32{,}51\%$$

	A	B	C	D
1	Cálculo das Probabilidades			
2				
3	C(26,2)	325	<--=COMBIN(26;2)	
4	C(26,3)	2600	<--=COMBIN(26;3)	
5	C(52,5)	2598960	<--=COMBIN(52;5)	

Assim a probabilidade de selecionar aleatoriamente 2 cartas vermelhas é 32,51%

EXEMPLO 3

Quando é feita amostragem de população finita sem reposição, a distribuição binomial não pode ser usada porque os eventos não são independentes. Daí então a distribuição hipergeométrica é usada. Esta é dada por

$$P_{hipergeométrica}=\frac{\binom{N-X_t}{n-X}\binom{X_t}{X}}{\binom{N}{n}}\quad \text{distribuição hipergeométrica}$$

Ela mede o número de sucessos X numa amostra de tamanho n extraída aleatoriamente e sem reposição de uma população de tamanho N, da qual X_t itens têm a característica de denotar sucesso.

a. Usando a fórmula, determine a probabilidade de extrair 2 homens numa amostra de 6 selecionada aleatoriamente sem reposição de um grupo de 10 pessoas, 5 das quais são homens.

b. Qual resultado teria sido se tivéssemos (incorretamente) usado a distribuição binomial?

Solução

a. Aqui X = 2 homens, n = 6, N = 10 e X_t = 5

$$P_{hipergeométrica}=\frac{\binom{10-5}{6-2}\binom{5}{2}}{\binom{10}{6}}=\frac{\binom{5}{4}\binom{5}{2}}{\binom{10}{6}}=\frac{\frac{5!}{4!\,1!}\frac{5!}{2!\,3!}}{\frac{10!}{6!\,4!}}=\frac{(5)(10)}{210}\cong 0{,}24$$

b. $P(2)=\frac{n!}{X!(n-X)!}p^X(1-p)^{n-X}=\frac{6!}{2!4!}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^4=\frac{15}{64}=0{,}23$

Seria notado que a amostra é muito pequena em relação à população (digamos, menos do que 5% da população), amostragem sem reposição tem pouco efeito na probabilidade de sucesso em cada tentativa e a distribuição binomial (que é mais fácil de usar) é uma boa aproximação para a distribuição hipergeométrica.

2. DISTRIBUIÇÃO HIPERGEOMÉTRICA NO EXCEL

O Excel dispõe da função estatística DIST.HIPERGEOM cuja sintaxe é:

DIST.HIPERGEOM(exemplo_s;exemplo_núm;população_s;num_população))

Esta função dá a probabilidade de acontecer um **número determinado de sucessos** na amostra exemplo_s, conhecidos o tamanho da amostra exemplo_núm, o número de sucessos na população população_s e o tamanho da população num_população. Por exemplo, a probabilidade de acontecerem 3 sucessos na amostra, conhecidos o tamanho da amostra 5, o número de sucessos na população 90 e o tamanho da população 500 é igual a **0,0386**, valor obtido com a fórmula: = DIST.HIPERGEOM(3;5;90;500) como mostra a planilha abaixo:

	A	B	C	D	E	F
13	**Função DIST.HIPERGEOM**					
14						
15		***x = nº de sucesso na amostra***	3			
16		***n = tamanho da amostra***	5			
17		***k=nº sucesso população***	90			
18		***N=tamanho população***	500			
19		***P(x)***	**0,0386**	<--=DIST.HIPERGEOM(C15;C16;C17;C18)		
20						

EXERCÍCIOS RESOLVIDOS

1. Suponha que selecionemos 5 cartas de baralho de um maço ordinário de jogo de baralho. Qual a probabilidade de obter 2 copas ou menos?

Solução

N = 52 k = 13 copas no maço n = 5 cartas selecionadas aleatoriamente X = 0 até 2

Liguemos estes valores na fórmula hipergeométrica como segue:

h(X≤x;N,n,k) = h(X≤2;52,5,13)= h(X=0;52,5,13) + h(X=1;52,5,13) + h(X=2;52,5,13)

$$h(X\leq 2;52,5,13) = [\frac{C_0^{13}C_{(5)}^{(52-13)}}{C_5^{52}}] + [\frac{C_1^{13}C_{(4)}^{(52-13)}}{C_5^{52}}] + [\frac{C_3^{13}C_{(3)}^{(52-13)}}{C_5^{52}}] = [\frac{1x\ 575.757}{2.598.960}] + [\frac{13x\ 82.251}{2.598.960}] + [\frac{78\ x\ 9.139}{2.598.960}]$$

	A	B	C	D
1	Cálculo das Probabilidades			
2				
3	C(13,0)	1	<--=COMBIN(13;0)	
4	C(39,5)	575757	<--=COMBIN(39;5)	
5	C(52,5)	2598960	<--=COMBIN(52;5)	
6	C(13,1)	13	<--=COMBIN(13;1)	
7	C(39,4)	82251	<--=COMBIN(39;4)	
8	C(13,2)	78	<--=COMBIN(13;2)	
9	C(39,3)	9139	<--=COMBIN(39;3)	

h(X≤2;52,5,13)= [0,221534] + [0,41142] + [0,27428]

h(X≤2;52,5,13)= 0,9072 ou 90,72%.

Assim a probabilidade de selecionar aleatoriamente no máximo 2 copas é 90,72%

EXERCÍCIOS

1. Determine a probabilidade de obtermos

Distribuição Multinomial

Um **experimento multinomial** é um experimento estatístico que tem as seguintes propriedades:

- O experimento consiste de n tentativas repetidas.
- Cada tentativa tem um número discreto resultados possíveis.
- Em qualquer tentativa dada, a probabilidade de que um particular resultado ocorrerá é constante.
- As tentativas são independentes; isto é, o resultado de uma tentativa não afeta o resultado das outras tentativas.

Considere o seguinte experimento estatístico. Você lança dois dados, três vezes e registra o resultado de cada lançamento. Este é um experimento multinomial, porque:

- O experimento consiste de tentativas repetidas. Lançamos o dado 3 vezes.
- Cada tentativa pode resultar num número discreto de resultados – 2 até 12.
- A probabilidade de qualquer resultado é constante; ela não muda de um lançamento para o próximo.
- As tentativas são independentes; isto é, obter um resultado particular numa tentativa não afeta o resultado das outras tentativas.

Nota: Um experimento binomial é um caso especial de um experimento multinomial. Aqui está a principal diferença. Com um experimento binomial, cada tentativa pode resultar em dois – e somente dois – resultados possíveis. Com um experimento multinomial, cada tentativa pode ter dois ou mais resultados possíveis.

Função de Probabilidade

Uma **distribuição multinomial** é a função distribuição de probabilidade dos resultados de um *experimento multinomial*. A fórmula multinomial define a probabilidade de qualquer resultado de um experimento multinomial.

Suponha um experimento multinomial que consiste de n tentativas, e cada tentativa pode resultar em quaisquer dos k resultados possíveis: $E_1, E_2, ..., E_k$. Suponha, além disso, que cada resultado possível possa ocorrer com probabilidades $p_1, p_2, p_3, ..., p_k$. Então a probabilidade p que E_1 ocorra n_1 vezes, E_2 ocorra n_2 vezes, ..., e E_k ocorra n_k vezes é:

$$P = \left[\frac{n!}{(n_1!\,.\,n_2!\,....\,n_k!)}\right].\left(p_1^{n_1}\,.\,p_2^{n_2}\,....\,p_k^{n_k}\right)$$

Onde $n = n_1 + n_2 + ... + n_k$.

Os exemplos abaixo ilustram como usar a fórmula multinomial para calcular a probabilidade de um resultado de um experimento multinomial.

EXEMPLO 1

Suponha uma carta de baralho sendo extraída aleatoriamente de um maço de jogo de baralho, e depois então devolvida ao maço. Este exercício é repetido 5 vezes. Qual é a probabilidade de se extraírem 1 espada, 1 copa, 1 ouros e 2 paus?

Solução:

Para resolver este problema, aplicamos a fórmula multinomial. Sabemos o seguinte:

- O experimento consiste de 5 tentativas, assim n = 5.
- As 5 tentativas produzem 1 espada, 1 copas, 1 ouros e 2 paus; assim $n_1 = 1$, $n_2 = 1$, $n_3 = 1$ e $n_4 = 2$
- Em qualquer tentativa particular, a probabilidade de extraírem 1 espada, cops, ouros ou paus é 0,25, 0,25, 0,25 e 0,25, respectivamente. Assim, $p_1 = 0,25$, $p_2 = 0,25$, $p_3 = 0,25$ e $p_4 = 0,25$

Liguemos estas entradas na fórmula multinomial, como mostrado abaixo:

$$P = \left[\frac{n!}{(n_1!\,.n_2!....n_k!)}\right].\left(p_1^{n_1}\,.\,p_2^{n_2}\,....\,p_k^{n_k}\right) = \left[\frac{5!}{(1!\,.1!\,.1!\,.2!)}\right].(0{,}25^1\,.\,0{,}25^1\,.\,0{,}25^1\,.\,0{,}25^2) = 0{,}05859$$

Assim, se extrairmos 5 cartas com reposição de um maço de cartas de baralho, a probabilidade de extrairmos 1 espada, 1 copa, 1 ouros e 2 paus é 0,05859 ou 5,859%.

EXEMPLO 2

Suponha que temos um vaso com 10 bolinhas de gude – 2 bolinhas vermelhas, 3 bolinhas verdes e 5 bolinhas azuis. Selecionamos 4 bolinhas aleatoriamente do vaso, **com reposição.** Qual é a probabilidade de selecionar 2 bolinhas verdes e 2 bolinhas azuis?

Solução:

Para resolver este problema, aplicamos a fórmula multinomial. Sabemos o seguinte:

- O experimento consiste de 4 tentativas, assim n = 4.
- As 4 tentativas produzem 0 bolinhas vermelhas, 2 bolinhas verdes e 2 bolinhas azuis; então $n_{vermelho} = 0$, $n_{verde} = 2$ e $n_{azul} = 2$.
- Em qualquer tentativa particular, a probabilidade de extraírem 1 vermelha, verde ou azul é 0,2, 0,3 e 0,5, respectivamente. Assim, $p_{vermelha} = 0,2$, $p_{verde} = 0,3$ e $p_{azul} = 0,5$.

Liguemos estas entradas na fórmula multinomial, como mostrado abaixo:

$$P = \left[\frac{n!}{(n_1!\,.n_2!\ldots.n_k!)}\right].\left(p_1^{n_1}.p_2^{n_2}\ldots.p_k^{n_k}\right) = \left[\frac{4!}{(0!\,.2!\,.2!)}\right].\left(0{,}2^0\,.\,0{,}3^2\,.\,0{,}5^2\right) = 0{,}135$$

Assim, se extrairmos 4 bolinhas **com reposição** de um vaso, a probabilidade de extrairmos 0 bolinhas vermelhas, 2 bolinhas verdes e 2 bolinhas azuis é 0,135 ou 13,5%