Ciência dos Dados

Modelos Probabilísticos Discretos

Distribuição de Bernoulli
Distribuição Binomial
Distribuição Hipergeométrica
Distribuição de Poisson

Magalhães e Lima, 7ª edição. Seções 3.2 e 3.3

Objetivos de Aprendizagem

Os alunos devem ser capazes de:

- Especificar as distribuições de probabilidades adequadas para variáveis aleatórias discretas considerando modelos probabilísticos discretos já bem definidos na literatura estatística.

Acompanhe, previamente, o PLANO DE AULA no BLACKBOARD!

Insper

Distribuição de Bernoulli

LEITURA COMPLEMENTAR

Contexto

Modelagem de fenômenos aleatórios que seguem padrões comuns.

Exemplos:

A1) Sortear uma amostra de 200 eleitores e contar quantos apoiam o governo.

A2) Observar, em uma amostra de 100 voos, quantos se atrasam.

B1) Observar o número de acessos a um site a cada hora.

B2) Registrar o número de reclamações recebidas por uma central telefônica em duas horas.

Insper

Ensaios de Bernoulli

O que há em comum nos seguintes eventos:

- oferecer uma apólice de seguro em um telefonema;
- observar a ocorrência de sinistro em um carro segurado por uma companhia de seguros;
- perguntar a um telespectador se ele se lembra de um anúncio comercial exibido num determinado horário.;
- verificar se um ativo se valorizou num determinado dia.
- observar se um resistor comprado na Santa Efigênia funciona perfeitamente.

Como definir uma v.a. para modelar esses fenômenos?

Ensaio de Bernoulli

Experimento que tem apenas dois resultados possíveis. Por convenção denomina-se o resultado de interesse como ***sucesso*** e o outro como ***fracasso***.

Jacob Bernoulli

1654-1705

Suiça

http://www-gap.dcs.st-and.ac.uk/~history/Mathematicians/Bernoulli_Jacob.html

Distribuição de Bernoulli

Assim, podemos definir uma variável aleatória X que assume dois valores:

$$X = \begin{cases} 1, & \text{se ocorrer sucesso} \\ 0, & \text{caso contrário} \end{cases}$$

Além disso, vamos considerar que a probabilidade de ocorrer o evento de interesse (sucesso) é *p*; e por consequência, a probabilidade de fracasso é $1-p$, sendo $0<p<1$.

Distribuição de Bernoulli

Utilizada na modelagem de ensaios de Bernoulli

$$X \sim Bern(p)$$

$$P(X=1) = p \quad \text{e} \quad P(X=0) = 1 - p$$

ou

$$\boxed{P(X=x) = p^x (1-p)^{1-x} \text{ ; } x=0 \text{ ou } x=1}$$

Determine E(X) e Var(X).

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

Distribuição Binomial

Vários ensaios de Bernoulli

Imagine, agora, que repetimos um ensaio de Bernoulli *n* vezes.

Suponha ainda que as repetições sejam independentes.

Uma amostra particular será constituída de uma sequência de sucessos e fracassos, ou, alternativamente, de uns e zeros.

Situações Práticas

Em muitas situações práticas, estamos interessados na probabilidade de **um evento ocorrer y vezes** em **n repetições do experimento**, por exemplo:

- ✓ a probabilidade de **vender 50 seguros** em **200 telefonemas**;
- ✓ a probabilidade de que **15 carros** de uma determinada marca sejam **roubados** na cidade, **dentre os 40 carros** desta marca segurados por uma companhia;
- ✓ a probabilidade de **100** em **300** telespectadores **entrevistados lembrarem quais produtos foram anunciados** em determinado programa;
- ✓ a probabilidade de **uma ação subir em 10** dos **21 dias avaliados**.

Experimento Binomial

- é uma sequência de ***n* repetições** (ou tentativas ou ensaios) idênticas;

- cada repetição tem apenas 2 resultados possíveis: um é denominado **sucesso** e o outro, **fracasso**;

- a **probabilidade de sucesso** para cada ensaio é denominada ***p*** e será constante em cada repetição. Então, a **probabilidade de fracasso** (***1-p***) também não varia de tentativa para tentativa;

- As tentativas são independentes.

Experimento Binomial – Exemplo 1

100K Ω

Um resistor de 100K Ohms comprado na Santa Efigênia tem probabilidade de falha de 20%, segundo um fabricante bem ruim.

Um aluno de engenharia compra um pacote com 3 desses resistores.

Responda:

a) Qual a probabilidade de exatamente dois falharem?

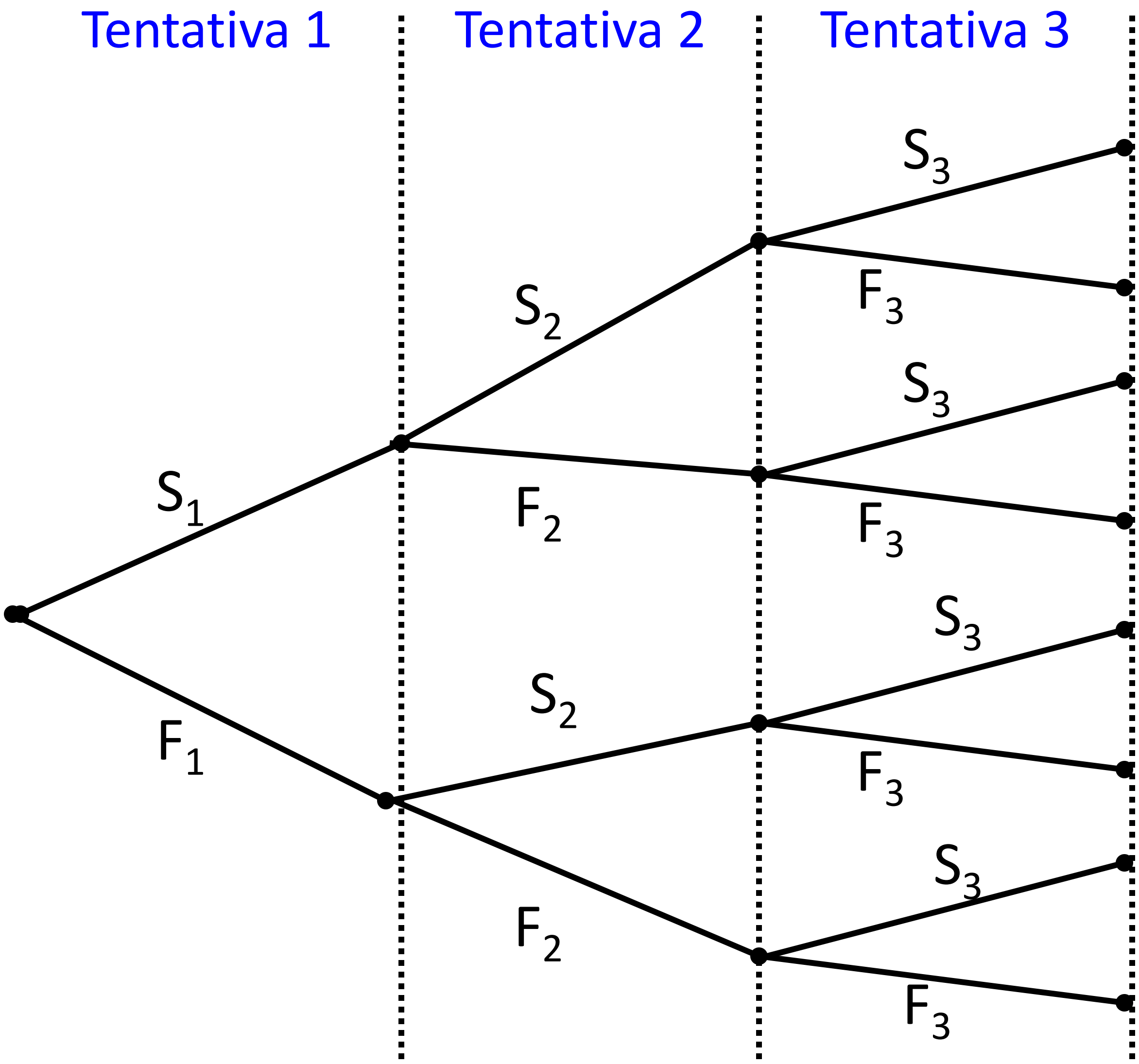
b) E se for um pacote com 100 resistores, qual a probabilidade de 20 falharem?

Exemplo 1

Tal experimento possui a distribuição binomial?

- Experimento consiste de 3 abordagens idênticas ➔ $n = 3$.
- Dois resultados possíveis em cada tentativa: falha (sucesso) ou não (fracasso).
- A probabilidade de falha é a mesma em cada tentativa ➔ $p = 0{,}2$
- As tentativas (resistores) são independentes.

Diagrama de Árvore



Resultados	Valor de Y
(S_1,S_2,S_3)	3
(S_1,S_2,F_3)	2
(S_1,F_2,S_3)	2
(S_1,F_2,F_3)	1
(F_1,S_2,S_3)	2
(F_1,S_2,F_3)	1
(F_1,F_2,S_3)	1
(F_1,F_2,F_3)	0

Distribuição de Probabilidades

Evento	y	$p(Y=y)$	Expressão	Geral
$S_1S_2S_3$	3	0,008	ppp	$p^3\,(1-p)^0$
$S_1S_2F_3$	2	0,032	$pp(1-p)$	$p^2\,(1-p)$
$S_1F_2S_3$	2	0,032	$p\,(1-p)\,p$	$p^2\,(1-p)$
$F_1S_2S_3$	2	0,032	$(1-p)\,pp$	$p^2\,(1-p)$
$S_1F_2F_3$	1	0,128	$p\,(1-p)\,(1-p)$	$p\,(1-p)^2$
$F_1S_2F_3$	1	0,128	$(1-p)\,p\,(1-p)$	$p\,(1-p)^2$
$F_1F_2S_3$	1	0,128	$(1-p)\,(1-p)\,p$	$p\,(1-p)^2$
$F_1F_2F_3$	0	0,512	$(1-p)\,(1-p)\,(1-p)$	$(1-p)^3$

$$P(Y=2) = 3\,p^2(1-p)$$

Considerando que p=0,2, temos:

$$P(Y=2) = 3 \times 0{,}2^2 \times (1-0{,}2) = 3 \times 0{,}032 = 0{,}096$$

Número de amostras gerando exatamente *y* sucessos em *n* tentativas

Amostras	*y*
$S_1S_2S_3$	3
$S_1S_2F_3$	**2**
$S_1F_2S_3$	**2**
$F_1S_2S_3$	**2**
$S_1F_2F_3$	1
$F_1S_2F_3$	1
$F_1F_2S_3$	1
$F_1F_2F_3$	0

Processo de Contagem

Combinatória

$$\boxed{C_{n,y} = \binom{n}{y} = \frac{n!}{y!(n-y)!}}$$

$$C_{3,2} = \binom{3}{2} = \frac{3!}{2!(3-2)!} = 3$$

Probabilidade para sequências com 2 sucessos

$p = 0,20$ = probabilidade de um resistor falhar

Resultado	Probabilidade	
$S_1S_2F_3$	$pp(1-p)$	$p^2(1-p)$
$S_1F_2S_3$	$p(1-p)p$	$p^2(1-p)$
$F_1S_2S_3$	$(1-p)pp$	$p^2(1-p)$

$$P(Y = 2) = 3\ p^2\ (1-p)$$

Probabilidade para sequências com 1 sucesso

p = 0,20 = probabilidade de um resistor falhar

Resultado	Probabilidade	
$S_1F_2F_3$	p(1-p)(1-p)	$p(1-p)^2$
$F_1S_2F_3$	(1-p)p(1-p)	$p(1-p)^2$
$F_1F_2S_3$	(1-p)(1-p)p	$p(1-p)^2$

$P(Y = 1) = 3\ p\ (1-p)^2$

Probabilidades para sequências com 0 e 3 sucessos

Resultado	Probabilidade	
$F_1F_2F_3$	(1-p)(1-p)(1-p)	$p^0(1-p)^3$

$$P(Y = 0) = (1-p)^3$$

Resultado	Probabilidade	
$S_1S_2S_3$	ppp	$p^3(1-p)^0$

$$P(Y = 3) = p^3$$

Distribuição de Probabilidades

Y: número de sucessos, com p = 0,20

Resultado	*y*	*P(Y=y)*
$S_1S_2S_3$	3	0,008
$S_1S_2F_3$	2	0,032
$S_1F_2S_3$	2	0,032
$F_1S_2S_3$	2	0,032
$S_1F_2F_3$	1	0,128
$F_1S_2F_3$	1	0,128
$F_1F_2S_3$	1	0,128
$F_1F_2F_3$	0	0,512
	soma	**1,000**

Distribuição de probabilidades de Y

y	*p(Y=y)*
0	0,512
1	0,384
2	0,096
3	0,008
soma	**1,000**

Exemplo 1

Um resistor de 100K Ohms comprado na Santa Efigênia tem probabilidade de falha de 20%, segundo um fabricante bem ruim.

Um aluno de engenharia compra um pacote com 3 desses resistor.

Responda:

a) Qual a probabilidade de exatamente dois falharem?

b) E se for um pacote com 100 resistores, qual a probabilidade de 20 falharem?

P(Z=20)

Z: número de resistores que falham em 100.
Evento de interesse (Z=20)
Sucesso: um resistor falhar
Probabilidade de um resistor falhar = 0,20

Pontos do espaço amostral	Resistor												Probab.
	1	2		19	20	21	22	...	81	...	99	100	
1	S	S	...	S	S	F	F	...	F	...	F	F	$0{,}2^{20}\ 0{,}8^{80}$
2	S	S	...	S	F	S	F	...	F	...	F	F	$0{,}2^{20}\ 0{,}8^{80}$
C	F	F	...	F	F	F	F	...	S	...	S	S	$0{,}2^{20}\ 0{,}8^{80}$

$$P(Z=20) = C\ 0{,}2^{20}\ 0{,}8^{80} = C\ 0{,}2^{20}\ (1 - 0{,}2)^{100-20}$$

Pontos com *20* sucessos em *100* tentativas

Clientes 1 2 3 ... 99 100

C= número de situações em que posso escolher a posição dos 20 resistores que falham.

$$\boxed{C = C_{100,20} = \binom{100}{20} = \frac{100!}{20!(100-20)!}}$$

$$P(Z=20) = \binom{100}{20} 0{,}2^{20}(1-0{,}2)^{100-20}$$

Distribuição Binomial

Modela experimentos binomiais.

Y: número de sucessos em um experimento binomial com n tentativas

Y ~ Bin(n;p)

Para y = 0,1,..., n, temos:

$$P(Y = y) = \binom{n}{y} p^{y} (1-p)^{n-y}$$

Em n repetições, em quantas combinações aparecem y sucessos?

Probabilidade de sucesso ocorrendo y vezes

Probabilidade de fracasso ocorrendo (n-y) vezes

Forma geral

$$P(Y = y) = \binom{n}{y} p^{y} (1 - p)^{n-y}$$

- P(Y=y): probabilidade de y sucessos em n tentativas
- n: número de tentativas
- p: probabilidade de sucesso em cada tentativa
- $\binom{n}{y} = \dfrac{n!}{y!(n-y)!}$

Esperança e Variância da Binomial

Quando **Y ~ Bin(n;p)**, então

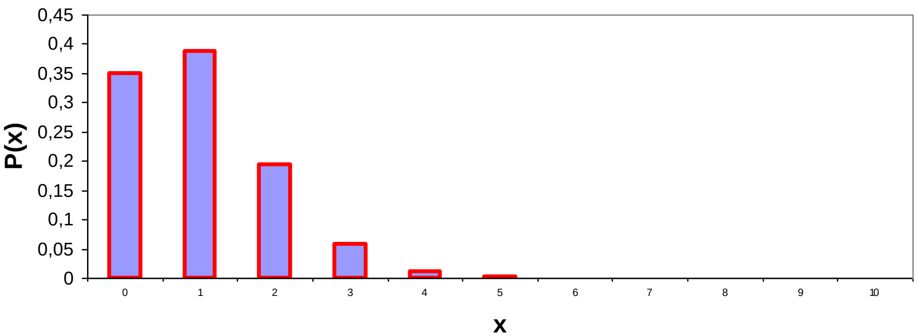
$$\mu = E(Y) = np$$

$$\sigma^2 = Var(Y) = np(1-p)$$

n = 10

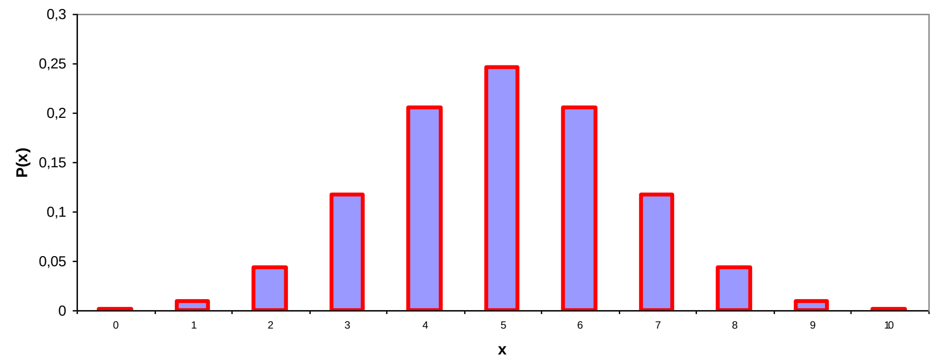
p=0,10

n=10, p=0,10



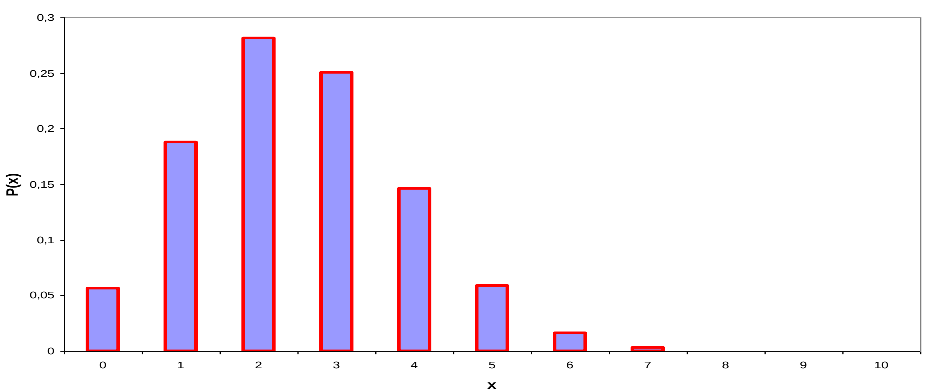
p=0,50

n=10, p=0,50



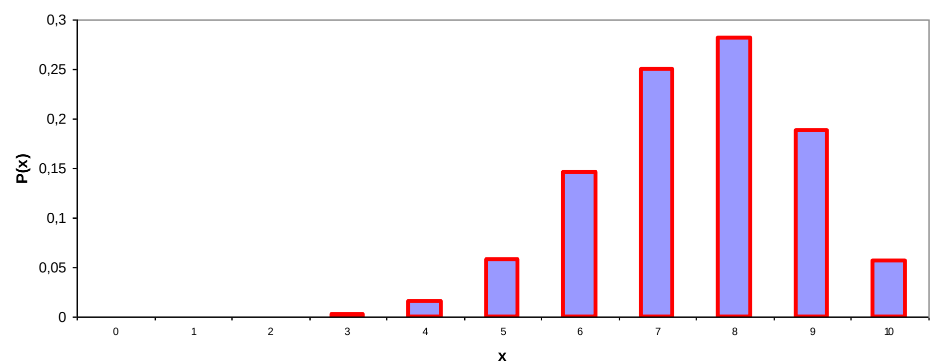
p=0,25

n=10, p=0,25



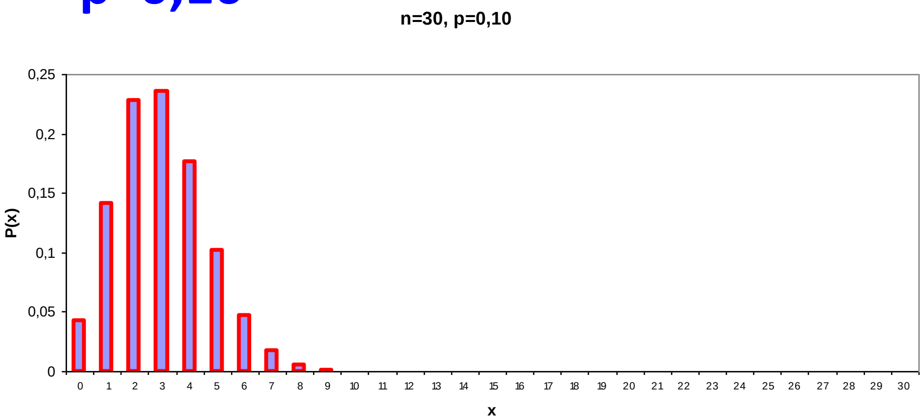
p=0,75

n=10, p=0,75

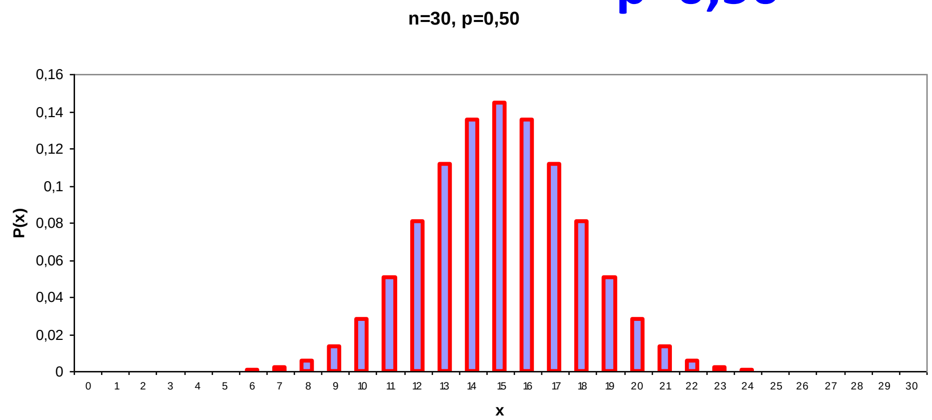


n = 30

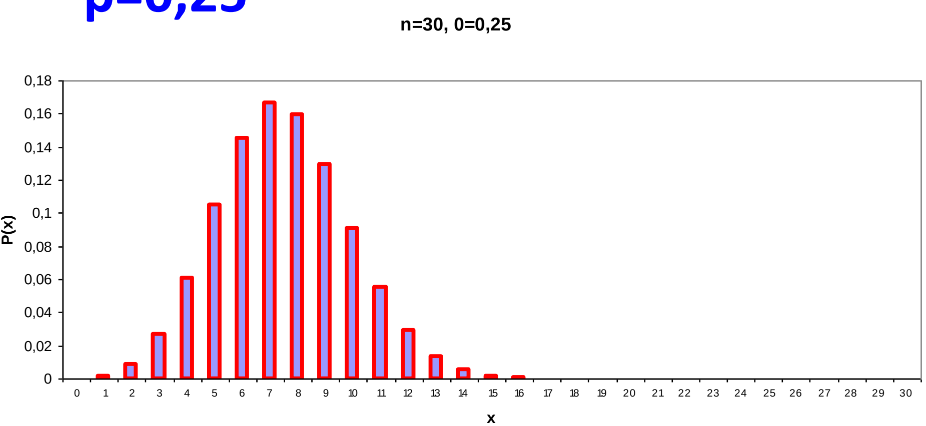
p=0,10



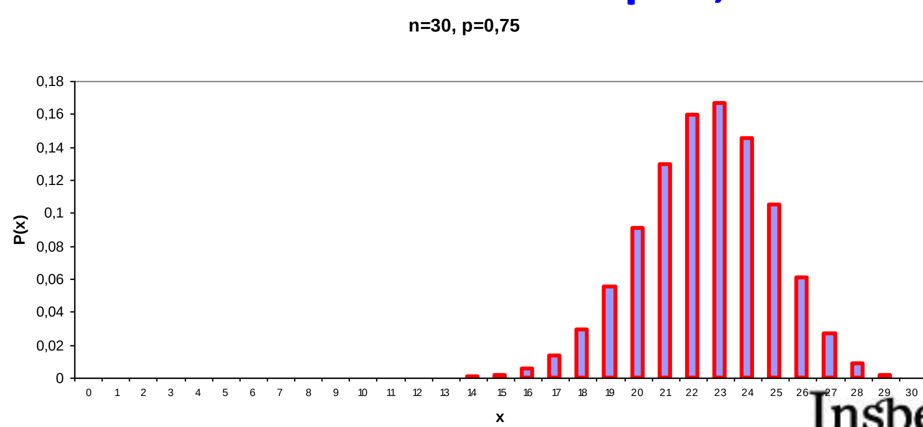
p=0,50



p=0,25



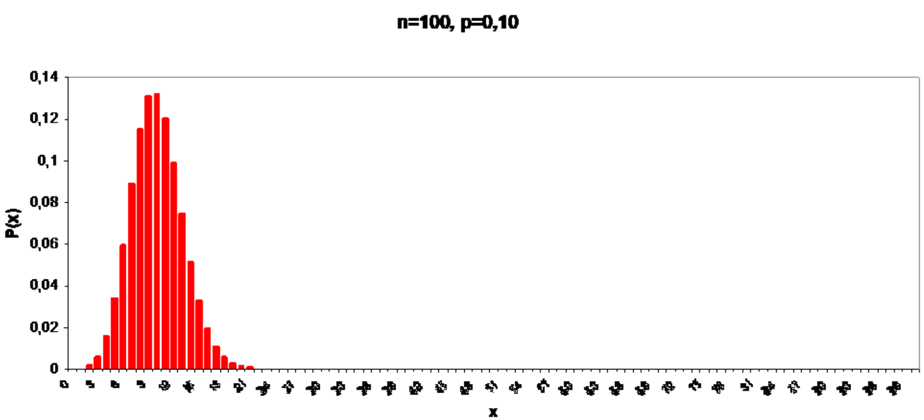
p=0,75



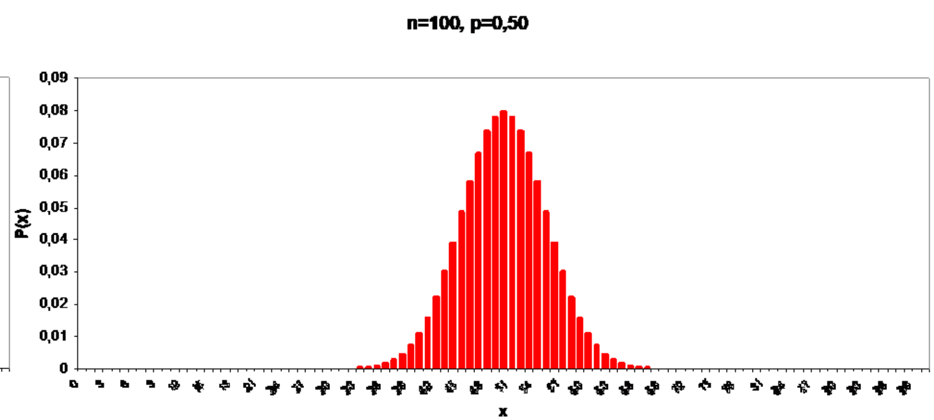
Insper

n = 100

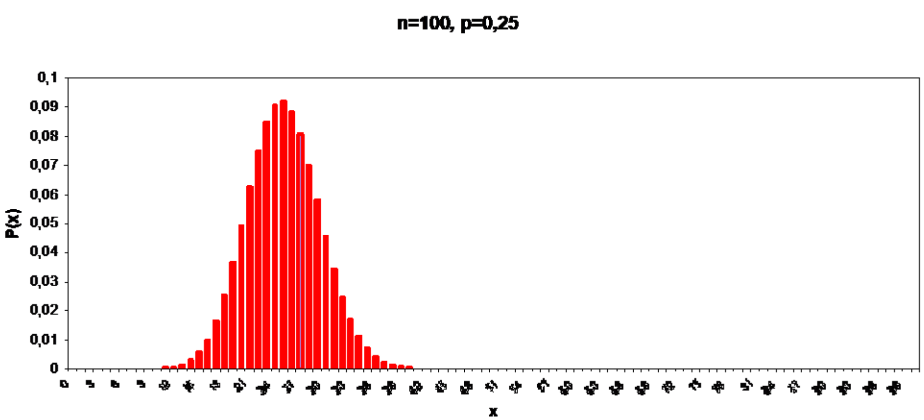
p=0,10



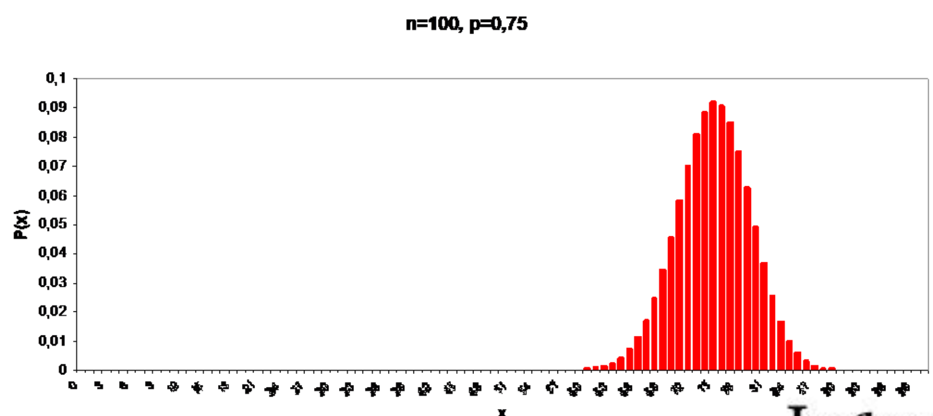
p=0,50



p=0,25



p=0,75



Experimento Binomial

A probabilidade de um determinado resistor falhar é sempre de 0,20.

Suponha que os resistores se comportem de maneira independente.

a) Se um pacote tem 100 resistores, qual a probabilidade de 20 falharem?

b) Se um pacote tem 100 resistores, qual a probabilidade de no máximo 20 falharem?

c) Se um corretor visita 100 clientes, qual a probabilidade de pelo menos 20 falharem?

Distribuição de Binomial

Como calcular no Python:

$$\mathbf{Y \sim Binomial(n, p)}$$

In [ ]: from scipy import stats

$$P(Y = y) \Rightarrow \text{stats.binom.pmf}(y, n, p)$$

$$P(Y \leq y) \Rightarrow \text{stats.binom.cdf}(y, n, p)$$

$$E(Y) \Rightarrow \text{stats.binom.mean}(n, p)$$

$$Var(Y) \Rightarrow \text{stats.binom.var}(n, p)$$

$$DP(Y) \Rightarrow \text{stats.binom.std}(n, p)$$

Exemplo 2

Determinado produto é vendido em caixas com 1000 peças. É uma característica da fabricação produzir 10% defeituosos. Normalmente, cada caixa é vendida por 12,00 unidades monetárias (u.m.).

Um comprador faz a seguinte proposta: de cada caixa, ele escolhe uma amostra de 20 peças (com reposição); se tiver 0 defeituoso, ele paga 18,00 u.m.; 1 ou 2 defeituosos, ele paga 12,00 u.m.; 3 ou mais defeituosos ele paga 6,00 u.m.. Qual alternativa é a mais vantajosa para o fabricante?

Admita independência entre as peças.

Solução:

- **Y: número de peças defeituosas dentre as 20 selecionadas**

$$Y \sim \mathrm{Bin}(20;\ 0{,}10)$$

$P(Y = 0) = 0{,}1216$

$P(Y = 1) = 0{,}2702$

$P(Y = 2) = 0{,}2852$

$P(Y \geq 3) = 1 - P(Y < 3) = 1 - 0{,}6770 = 0{,}3230$

Solução:

W: valor pago pela caixa segundo a proposta do comprador

w	P(W=w)
6	0,3230
12	0,5554 (0,2702+0,2852)
18	0,1216

E(W) =18x0,1216 + 12x(0,2702+0,2852) + 6x0,3230 = 10,79

Conclusão: A proposta do comprador é menos vantajosa para o fabricante.

Distribuição Hipergeométrica

LEITURA COMPLEMENTAR

Exemplo 3
versão simples

De 50 investimentos gerenciados por um banco, 15 tiveram rentabilidade positiva no ano passado.

Se um cliente sortear uma amostra ***sem reposição*** de **3 investimentos** gerenciados pelo banco, qual é a **probabilidade de exatamente 2 terem tido rentabilidade positiva**?

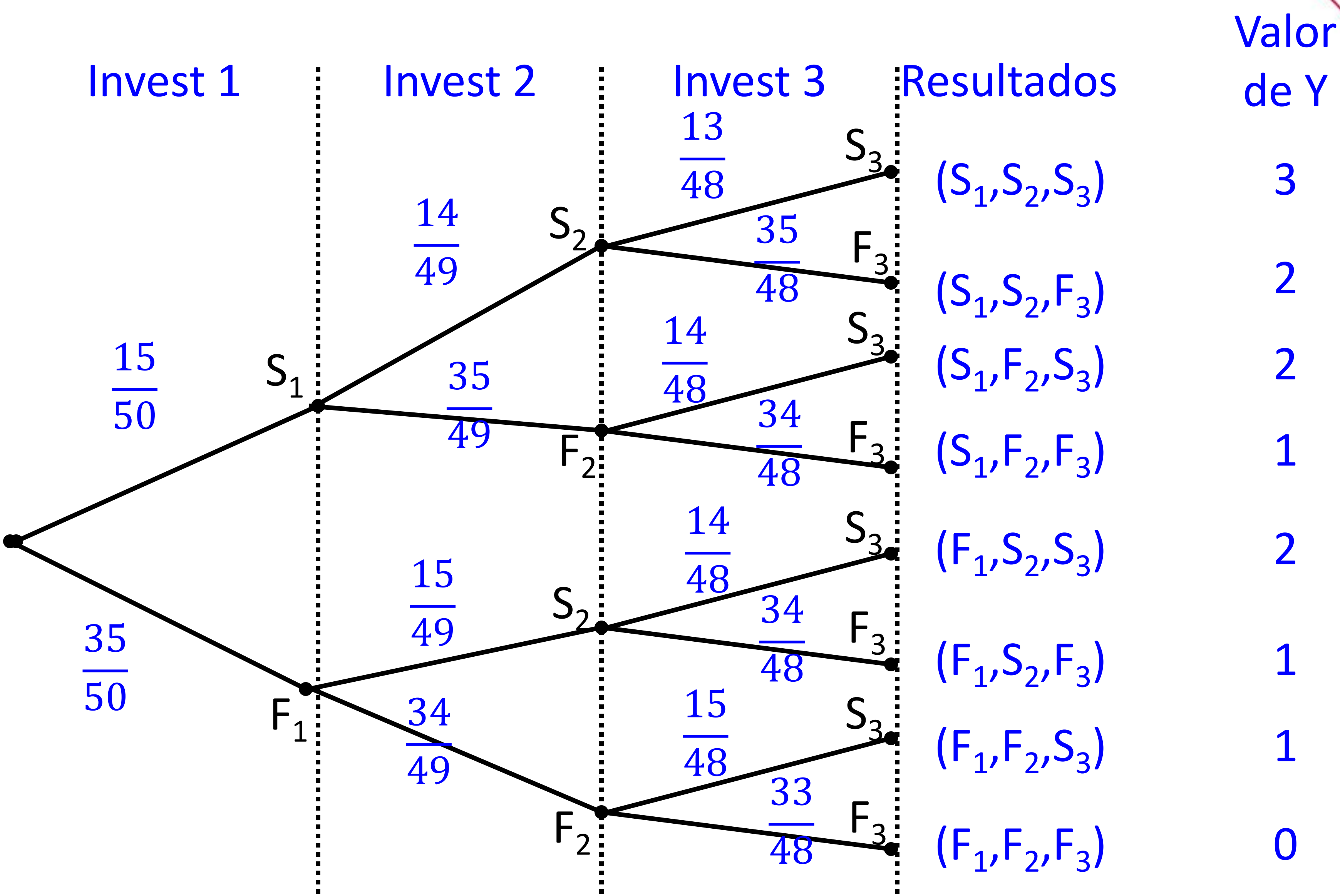
Exemplo 3 – versão simples

Tal experimento possui a distribuição binomial?

- Experimento **NÃO** consiste de 3 abordagens idênticas ➔ n = 3.
- Dois resultados possíveis em cada tentativa: rentabilidade positiva (sucesso) ou rentabilidade não positiva (fracasso).
- A probabilidade de rentabilidade positiva muda de acordo com o resultado da tentativa anterior.
- As tentativas **NÃO** são independentes.

NÃO POSSUI DISTRIBUIÇÃO BINOMIAL!!!

Diagrama de Árvore



Insper

Distribuição de Probabilidades

Amostras	y
(S_1,S_2,S_3)	3
(S_1,S_2,F_3)	2
(S_1,F_2,S_3)	2
(S_1,F_2,F_3)	1
(F_1,S_2,S_3)	2
(F_1,S_2,F_3)	1
(F_1,F_2,S_3)	1
(F_1,F_2,F_3)	0

y	P(Y=y)
0	$\frac{35}{50} \bullet \frac{34}{49} \bullet \frac{33}{48} = 0{,}3339$
1	$3 \bullet \frac{35}{50} \bullet \frac{34}{49} \bullet \frac{15}{48} = 0{,}4554$
2	$3 \bullet \frac{35}{50} \bullet \frac{15}{49} \bullet \frac{14}{48} = 0{,}1875$
3	$\frac{15}{50} \bullet \frac{14}{49} \bullet \frac{13}{48} = 0{,}0232$
	1,000

Exemplo 3 – versão simples

$$P(Y = 2) = 3 \bullet \frac{35}{50} \bullet \frac{15}{49} \bullet \frac{14}{48} = \frac{\binom{15}{2}\binom{35}{1}}{\binom{50}{3}}$$

Distribuição Hipergeométrica

A distribuição hipergeométrica é parecida com a binomial só que as replicações não são independentes.

- ***r*** **= número de sucessos numa população de tamanho** ***N*****.**
- ***N* - *r*** **= número de fracassos nesta população.**
- ***n*** **= tamanho da amostra sem reposição.**
- **y = número de sucessos na amostra.**
- ***n*** **- y = número de fracassos na amostra.**

Distribuição Hipergeométrica

Número de maneiras que **n-y** fracassos podem ser selecionados em **N-r** fracassos na população

Número de maneiras que **y** sucessos podem ser selecionados em **r** sucessos na população

$$P(Y=y)=\frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}}$$

Número de maneiras que uma amostra de tamanho **n** pode ser selecionada de uma população de tamanho **N**

Distribuição Hipergeométrica

Y ~ hip(N,r,n)

$$P(Y = y) = \frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}} \quad y = \max\{0, n-(N-r)\}, \; \ldots, \; \min\{r,n\}$$

em que

$n =$ número de replicações;

$y =$ número de sucessos nas n replicações;

$r =$ número de sucessos na população.

Como calcular a Hipergeométrica usando Excel

Y ~ hip(N,r,n)

$$P(Y=y)=\frac{\binom{r}{y}\binom{N-r}{n-y}}{\binom{N}{n}}=dist.hipergeom.n(y;n;r;N;0)$$

$$P(Y\le y)=dist.hipergeom.n(y;n;r;N;1)$$

Distribuição Hipergeométrica

Y ~ hip(N,r,n)

Temos $$E(Y) = n\frac{r}{N} = np$$

$$Var(Y) = np(1-p)\frac{N-n}{N-1}$$

sendo $p = \dfrac{r}{N}$ a probabilidade de

sucesso no início do experimento.

Exemplo 3
versão + complexa

De 50 investimentos gerenciados por um banco, 15 tiveram rentabilidade positiva no ano passado.

Se um cliente sortear uma amostra ***sem reposição*** de **10 investimentos** gerenciados pelo banco, qual é a **probabilidade de exatamente 2 terem tido rentabilidade positiva**?

Exemplo 3
versão + complexa

N = 50 → investimentos gerenciados pelo banco

r = 15 → tiveram rentabilidade positiva

n = 10 → amostra ***sem reposição***

$$P(Y=2)=\frac{\binom{15}{2}\binom{35}{8}}{\binom{50}{10}}=0{,}2406$$

Relação entre as Distribuições Binomial e Hipergeométrica

Refaça o Exemplo 2 considerando extrações **sem reposição** Compare os resultados com e sem reposição

Resp.:

Com repos.: 10,79

Sem repos.: 10,78

Determinado produto é vendido em caixas com 1000 peças. É uma característica da fabricação produzir 10% defeituosos. Normalmente, cada caixa é vendida por 12,00 unidades monetárias (u.m.).

Um comprador faz a seguinte proposta: de cada caixa, ele escolhe uma amostra de 20 peças; se tiver 0 defeituoso, ele paga 18,00 u.m.; 1 ou 2 defeituosos, ele paga 12,00 u.m.; 3 ou mais defeituosos ele paga 6,00 u.m.. Qual alternativa é a mais vantajosa para o fabricante?

Solução: COM reposição

- **Y: número de peças defeituosas dentre as 20 selecionadas**

$$Y \sim \text{Bin}(20;\ 0{,}10)$$

$P(Y = 0) = 0{,}1216$

$P(Y = 1) = 0{,}2702$

$P(Y = 2) = 0{,}2852$

$P(Y \geq 3) = 1 - P(Y < 3) = 1 - 0{,}6770 = 0{,}3230$

- **Proposta do comprador:**

$$18 \times 0{,}1216 + 12 \times (0{,}2702 + 0{,}2852) + 6 \times 0{,}3230 = 10{,}79$$

- **Conclusão: A proposta do comprador é menos vantajosa para o fabricante.**

Solução: SEM reposição

- **Y: número de peças defeituosas dentre as 20 selecionadas**

$$Y \sim hip(1000;100;20)$$

$P(Y = 0) = 0{,}1190$

$P(Y = 1) = 0{,}2702$

$P(Y = 2) = 0{,}2881$

$P(Y \geq 3) = 1 - P(Y < 3) = 1 - 0{,}6773 = 0{,}3227$

- **Proposta do comprador:**

$$18 \times P(Y=0) + 12 \times [P(Y=1)+P(Y=2)] + 6 \times P(Y \geq 3) = 10{,}78$$

- **Conclusão: A proposta do comprador é menos vantajosa para o fabricante.**

Relação entre as Distribuições Binomial e Hipergeométrica

Propriedade:

Se N for grande, quando comparado com n, então extrações com ou sem reposição serão praticamente equivalentes, de modo que as probabilidades dadas pela distribuição exata serão aproximadamente iguais às fornecidas pela distribuição binomial.

O mesmo vale para o cálculo da esperança e variância da v.a. (note que $N - n \cong N - 1$, se $n << N$).

Distribuição de Poisson

Exemplo 4

Uma fábrica produz 24 horas por dia. Interrupções na linha de montagem por falha humana, ocorrem a uma taxa média de 4 a cada 24 horas.

a. Qual é a probabilidade de ocorrerem exatamente 3 interrupções em um dia de trabalho?

b. Qual é a probabilidade de ocorrem exatamente 3 interrupções em 2 dias de trabalho?

Experimento de Poisson

- Estimativa do número de ocorrências num espaço contínuo (tempo, área, distância, etc.).

Exemplo:

Número de aviões que decolam numa hora
Número de usuários conectados no Facebook por minuto
Número de acidentes em 100 km de uma rodovia
Número de clientes na fila por minuto (teoria das filas)
Consumo de um produto em um mês (gestão de estoques)

Propriedades (suposições):

- ✓ Probabilidade de uma ocorrência é a mesma para dois intervalos quaisquer de igual tamanho.
- ✓ A ocorrência ou não num dado intervalo é independente da ocorrência ou não em outro intervalo.

Siméon Denis Poisson

1781-1840

França

http://www-gap.dcs.st-and.ac.uk/~history/Mathematicians/Poisson.html

Distribuição de Poisson

Seja **X** o número de ocorrências de um evento num intervalo de tempo ou espaço, então

$$\mathbf{X \sim Poisson(\lambda)} \;\; \Rightarrow \;\; \mathbf{P(X = \mathit{x}) = \frac{\mathit{e}^{-\lambda}\,\lambda^{\mathit{x}}}{\mathit{x}!}},$$

em que

$x = 0, 1, 2, 3, \ldots$ e

λ é o número médio de eventos ocorrendo no intervalo considerado.

Voltando ao Exemplo 4

Uma fábrica produz 24 horas por dia. Interrupções na linha de montagem por falha humana, ocorrem a uma taxa média de 4 a cada 24 horas.

a. Qual é a probabilidade de ocorrerem exatamente 3 interrupções em um dia de trabalho? $P(X=3)=19{,}54\%$

b. Qual é a probabilidade de ocorrem exatamente 3 interrupções em 2 dias de trabalho? $P(Y=3)=2{,}86\%$

Distribuição de Poisson

Quando $\mathbf{X}$ tem distribuição de Poisson com taxa λ, ou seja,

$$\mathbf{X \sim Poisson(\lambda)}$$

com $x = 0, 1, 2, 3, \ldots$

Então, prova-se que

$$\mathbf{E(X) = Var(X) = \lambda}$$

Distribuição de Poisson

Como calcular no Python:

X~Poisson(λ)

In [ ]: from scipy import stats

$$P(X = x) \Rightarrow \text{stats.poisson.pmf}(x, \lambda)$$

$$P(X \leq x) \Rightarrow \text{stats.poisson.cdf}(x, \lambda)$$

$$E(X) \Rightarrow \text{stats.poisson.mean}(\lambda, \text{loc} = 0)$$

$$\text{Var}(X) \Rightarrow \text{stats.poisson.var}(\lambda, \text{loc} = 0)$$

$$\text{DP}(X) \Rightarrow \text{stats.poisson.std}(\lambda, \text{loc} = 0)$$

Exemplo 5

Uma central telefônica recebe uma média de 5 chamadas por minuto.

☆ Probabilidade de uma chamada é a mesma para dois intervalos de tempo de igual tamanho e a chegada ou não de uma chamada num dado instante de tempo é independente da chegada ou não de um chamada em outro intervalo de tempo.

☆ Seja X a v.a. que denota o número de chamadas por minuto, então

$$X \sim \text{Poisson}\ (5)$$

Exemplo 5 (cont.)

a) Qual a probabilidade de não ter chamada no próximo minuto?

b) Qual a probabilidade de chegar no máximo 2 chamadas no próximo minuto?

c) Qual a probabilidade de chegar 10 chamadas num período de 3 minutos?

Exemplo 5 (respostas)

a) Qual a probabilidade de não ter chamada no próximo minuto?

$$P(X = 0) = \frac{5^0 e^{-5}}{0!} = 0,0067 = 0,67\%$$

b) Qual a probabilidade de chegar no máximo 2 chamadas no próximo minuto?

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) =$$

$$= \frac{5^0 e^{-5}}{0!} + \frac{5^1 e^{-5}}{1!} + \frac{5^2 e^{-5}}{2!} =$$

$$= 0,0067 + 0,0337 + 0,084 = 0,1244$$

$$= 12,44\%$$

Exemplo 5 (respostas)

c) Qual a probabilidade de chegar 10 chamadas num período de <u>3 minutos</u>?

$$P(Y = 10) = \frac{15^{10} e^{-15}}{10!} = 0{,}0486 = 4{,}86\%$$

Observe que chegam, em média, 5 chamadas por minuto então num período de 3 minutos chegam em média 15 chamadas. Logo, temos uma nova variável Poisson com parâmetro 15.

Isto é, Y ~ Poisson (15)

Exercícios

Exercício 1

A mortalidade de microempresas no 1º ano de funcionamento é da ordem de 55%. Deseja-se avaliar as causas da mortalidade. Numa amostra de 10 empresas,

Resp.:

0,234 a. qual á a probabilidade de exatamente 5 terem falido ao final do 1º ano?

0,973 b. qual é a probabilidade de pelo menos 3 virem a falir no 1º ano?

0,481 c. sabe-se que pelo menos 3 faliram no 1º ano, qual é a probabilidade de no máximo 5 terem falido?

5,5

1,57 d. qual o número esperado de empresas que irão à falência no 1º ano na amostra? E o desvio padrão?

Exercício 2

O número médio de clientes satisfeitos com o atendimento de uma loja, em amostras de 40 clientes escolhidos ao acaso, é de 5,5.

Qual é a probabilidade de, numa amostra de 40 clientes escolhidos ao acaso, encontrarmos pelo menos 2 satisfeitos?

Resp.: p=0,1375 → P(X>=2)=98,02%

Exercício 3

Para estudar as causas de falência de uma microempresa em uma cidade', conta-se com o cadastro de 200 microempresas abertas no ano de 2000.

Resp.:

Sem repos.: 3,5%

Com repos.: 4,2%

a. Admitindo que 100 empresas tenham falido, qual é a probabilidade de se encontrar exatamente 20 empresas falidas numa amostra, sem reposição, de 50 empresas do cadastro? E numa amostra com reposição?

n=14

b. Numa amostragem sem reposição, qual deve ser o tamanho da amostra de modo que com no mínimo 90% de probabilidade tenhamos pelo menos 5 empresas falidas?

Insper

Exercício 4

Um supermercado que tem um cartão de fidelidade classifica os clientes em regular e diamante. Atualmente 70 % dos clientes são regulares. A direção do supermercado deseja premiar os clientes que, num determinado mês, comprem mais de 1 mil reais. Sabemos que se o cliente é diamante a probabilidade dele fazer uma compra acima de 1 mil reais é de 75%, já se o cliente regular é 5%.

Resp.:

6,09% a) Sorteada uma amostra de 20 clientes diamante, qual é a probabilidade de exatamente 12 fazerem uma compra acima de 1 mil reais?

1,83% b) Sorteados 20 clientes com a mesma classificação ao acaso, qual a probabilidade de exatamente 12 fazerem uma compra acima de 1 mil reais?

0,11% c) Sorteados 20 clientes ao acaso, qual a probabilidade de exatamente 12 fazerem uma compra acima de 1 mil reais?

7,05% d) Um gerente está analisando 200 compras efetuadas em um caixa, dessas 120 foram superiores a 1 mil reais. Dessas compras sob análise, deseja-se sortear uma amostra de 20 compras, sendo que a mesma compra não pode ser sorteada mais de uma vez, qual é a probabilidade de 15 terem sido superiores a 1 mil reais?

Exercício 5

Um pendrive permite 60 minutos de gravação de um vídeo em qualidade full HD. Sabe-se que defeitos de fabricação surgem a uma taxa de 1,2 a cada 10 horas de gravação.

a. Qual é a probabilidade de um pendrive apresentar defeito de fabricação? **11,31%**

b. Sabendo que um pendrive apresentou defeito, qual a probabilidade de que ele tenha no máximo dois defeitos? **99,77%**

c. Uma pessoa compra uma embalagem com 10 pendrives. Qual é a probabilidade de pelo menos 3 apresentarem defeito de fabricação? **9,4%**

d. Quais suposições você fez para responder os itens (a) e (c)? **Slide 3 – Aula 14 e Slide 9 - Aula 13, respectivamente.**

Exercício 6

Uma das responsabilidades da área de Tecnologia da Informação (TI) de uma empresa é controlar o fluxo de mensagens eletrônicas e detectar a presença de SPAMs (mensagens não desejadas).

Os responsáveis pela área de TI de uma empresa detectaram que um funcionário recebe no seu e-mail individual da empresa, em média, 3 SPAMs a cada quatro dias. Assuma que os SPAMs são enviados de forma independente para cada funcionário da empresa.

Resp.:

52,8% a) Qual a probabilidade de um funcionário receber SPAM em um dia?

15,14% b) Considere que a empresa tem, atualmente, 85 funcionários. Selecionando 40 funcionários aleatoriamente <u>(sem reposição)</u>, qual é a probabilidade de vinte deles receberem SPAM no seu e-mail individual em um dia qualquer?

11,82% c) Considere que a empresa tem, atualmente, 8.500 funcionários. Selecionando 40 funcionários aleatoriamente <u>(sem reposição)</u>, qual é a probabilidade de vinte deles receberem SPAM no seu e-mail individual em um dia qualquer?

11,79% d) Refaça item (c) assumindo sorteio com reposição!! Esse resultado era esperado?

Preparo para próxima aula

Os alunos devem se preparar com:

1. Leitura prévia necessária: Magalhães e Lima (7ª. Edição): Capítulo 6 - Seção 6.1.
2. Python.