Estatística

5 - Distribuição de Probabilidade de Variáveis Aleatórias Discretas

Principais Distribuições de Probabilidades

- Equiprovável
- Bernoulli
- Binomial
- Poisson
- Geométrica
- Pascal
- Hipergeométrica
- Multinomial

V. A. Discretas

- Uniforme
- Exponencial
- Normal
- Beta
- Log Normal
- Gama
- Weibull

V. A. Contínuas

Distribuição Equiprovável

Todos os possíveis valores da Variável Aleatória tem a mesma Probabilidade de ocorrer

n valores $\Rightarrow$ $\Pr(X = x_i) = \dfrac{1}{n}$

Para valores equi-espaçados (a diferença entre os valores é constante e igual a h), tem-se:

$$\boxed{E(X) = \frac{x_1 + x_n}{2}}$$

$$\boxed{Var(X) = \frac{h^2(n^2 - 1)}{12}}$$

Distribuição de Bernoulli

$$\text{Experimento} \begin{cases} \nearrow \text{“sucesso”} \\ \searrow \text{“fracasso”} \end{cases}$$

Seja X: variável aleatória com possíveis resultados:

X = 1 se o resultado for um sucesso

X = 0 se o resultado for um fracasso

p: probabilidade de ocorrer sucesso

q: probabilidade de não ocorrer sucesso (fracasso)

$$Pr(X) = \begin{cases} q = 1 - p & \text{para } X = 0; \\ p & \text{para } X = 1; \\ 0 & \text{para } X \neq 0 \text{ ou } X \neq 1 \end{cases}$$

$$\boxed{E(X) = p} \qquad \boxed{Var(X) = p.q}$$

Distribuição Binomial

Condições do experimento:

(1) número fixo de repetições independentes : n

(2) cada repetição tem Distribuição Bernoulli:

"sucesso" ← ou → "fracasso"

(3) Probabilidade *p* de sucesso é constante

Seja:

X: variável aleatória Binomial

n: número de repetições

k: número de sucessos

Pr(X=k): Probabilidade de k "sucessos" em n repetições

$$\Pr(X = k) = \binom{n}{k} p^k q^{n-k}$$

Onde:

$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$$

Distribuição Binomial

Pr(Y=k): Probabilidade de k "sucessos" nas primeiras k repetições de um total de n repetições

$$\underbrace{1, 1, 1, 1, 1, \ldots, 1}_{k} \qquad \underbrace{0, 0, 0, 0, \ldots, 0}_{n-k}$$

$$P(Y=k) = p^k . q^{n-k}$$

Considerando todas as combinações de n elementos k a k tem-se:

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

Obs.: Valores tabelados para n=10 e n=20

$$\boxed{E(x) = n.p} \qquad \boxed{Var(x) = n.p.q}$$

Distribuição Binomial

Exemplo: Lançamento de 4 moedas viciadas.
Probabilidade de sair cara (k) é 0,8 e coroa (c) é 0,2
Seja X: número de caras Logo: p=0,8 e q=0,2
Calcular a probabilidade de sair 2 caras: Pr(X=2)=?

1° modo:

$$S = \left\{ \begin{array}{ccccc} cccc & ccck & cckk & kkkc & kkkk \\ & cckc & ckck & kkck & \\ & ckcc & kcck & kckk & \\ & kccc & kkcc & ckkk & \\ & & kckc & & \\ & & ckkc & & \end{array} \right\}$$

Pr(kkcc) = ppqq = (0,8)(0,8)(0,2)(0,2) = 0,0256
Pr(kckc) = pqpq = (0,8)(0,2)(0,8)(0,2) = 0,0256
Pr(kcck) = pqpq = (0,8)(0,2)(0,2)(0,8) = 0,0256
Pr(ckkc) = pqpq = (0,2)(0,8)(0,8)(0,2) = 0,0256
Pr(ckck) = pqpq = (0,2)(0,8)(0,2)(0,8) = 0,0256
Pr(cckk) = pqpq = (0,2)(0,2)(0,8)(0,8) = 0,0256
Para esses 6 casos, tem-se:
Pr(X=2) = 6 (0,0256) = 0,1536

<u>Obs</u>: **Não usar a "regra"** = número de eventos favoráveis (6) / número de eventos possíveis (16)= 6/16=0,375=6x$(0,5)^4$, pois os eventos (sair cara, sair coroa) não são equiprováveis; **as moedas são viciadas !!**

Distribuição Binomial

2° modo:

Deseja-se calcular a probabilidade de sair 2 caras, em 4 jogadas da moeda

Considerando-se

sucesso: sair cara

n = 4

p = 0,8

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

$$\Pr(X = 2) = \binom{4}{2} \cdot p^2 \cdot q^{(4-2)}$$

$$\Pr(X = 2) = \frac{4!}{(4-2)!2!} \cdot 0{,}8^2 \cdot 0{,}2^{(4-2)}$$

$$\Pr(X = 2) = 0{,}1536$$

Tabela . Distribuição binomial

$n = 10$									
k \ p	0,02	0,05	0,10	0,20	0,25	0,30	0,40	0,50	
0	0,8171	0,5987	0,3487	0,1074	0,0563	0,0282	0,0060	0,0010	10
1	0,1667	0,3151	0,3874	0,2684	0,1877	0,1211	0,0403	0,0098	9
2	0,0153	0,0746	0,1937	0,3020	0,2816	0,2335	0,1209	0,0439	8
3	0,0008	0,0105	0,0574	0,2013	0,2503	0,2668	0,2150	0,1172	7
4	0,0000	0,0010	0,0112	0,0881	0,1460	0,2001	0,2508	0,2051	6
5		0,0001	0,0015	0,0264	0,0584	0,1029	0,2007	0,2461	5
6		0,0000	0,0001	0,0055	0,0162	0,0368	0,1115	0,2051	4
7			0,0000	0,0008	0,0031	0,0090	0,0425	0,1172	3
8				0,0001	0,0004	0,0014	0,0106	0,0439	2
9				0,0000	0,0000	0,0001	0,0016	0,0098	1
10						0,0000	0,0001	0,0010	0
$n = 20$									
0	0,6676	0,3585	0,1216	0,0115	0,0032	0,0008	0,0000		20
1	0,2725	0,3774	0,2702	0,0577	0,0211	0,0068	0,0005	0,0000	19
2	0,0528	0,1887	0,2852	0,1369	0,0669	0,0278	0,0031	0,0002	18
3	0,0065	0,0596	0,1901	0,2054	0,1339	0,0716	0,0123	0,0011	17
4	0,0006	0,0133	0,0898	0,2182	0,1897	0,1304	0,0350	0,0046	16
5	0,0000	0,0022	0,0319	0,1746	0,2023	0,1789	0,0746	0,0148	15
6		0,0003	0,0089	0,1091	0,1686	0,1916	0,1244	0,0370	14
7		0,0000	0,0020	0,0545	0,1124	0,1643	0,1659	0,0739	13
8			0,0004	0,0222	0,0609	0,1144	0,1797	0,1201	12
9			0,0001	0,0074	0,0271	0,0654	0,1597	0,1602	11
10			0,0000	0,0020	0,0099	0,0308	0,1171	0,1762	10
11				0,0005	0,0030	0,0120	0,0710	0,1602	9
12				0,0001	0,0008	0,0039	0,0355	0,1201	8
13				0,0000	0,0002	0,0010	0,0146	0,0739	7
14					0,0000	0,0002	0,0048	0,0370	6
15						0,0000	0,0013	0,0148	5
16							0,0003	0,0046	4
17							0,0000	0,0011	3
18								0,0002	2
19								0,0000	1
20									0
	0,98	0,95	0,90	0,80	0,75	0,70	0,60	0,50	k / p

Distribuição Binomial

Exemplo: Lançamento de 10 moedas viciadas.
Probabilidade de sair cara (k) é 0,8.
Deseja-se calcular a probabilidade de sair 6 caras

n = 10
p = 0,8
k = 6

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

$$\Pr(X = 6) = \binom{10}{6} \cdot 0{,}8^6 \cdot 0{,}2^{(10-6)}$$

$$\Pr(X = 6) = \frac{10!}{(10-6)!\,6!} \cdot 0{,}8^6 \cdot 0{,}2^4$$

$$\Pr(X = 6) = 210 \cdot 0{,}262 \cdot 0{,}0016$$

$$\Pr(X = 6) = 0{,}0881$$

Da Tabela, utilizando-se:
n=10 p=0,8 (rodapé da tabela) k=4

obtem-se: P(X=4) = 0,0881

Distribuição de Poisson

X: Número de sucessos em um determinado intervalo contínuo (tempo, comprimento, superfície, volume, etc).

Exemplos:

- ✓ Número de pessoas que chegam na rodoviária no período de 1 h.
- ✓ Número de defeitos em barras de aço 5 m.
- ✓ Número de focos de incêndio por hectare.

Hipóteses:

1. O número de sucessos em intervalos não sobrepostos constituem variáveis aleatórias independentes.
2. A probabilidade do número de sucessos em qualquer intervalo depende apenas da sua dimensão. Por outras palavras, em intervalos de mesma dimensão são iguais as probabilidades de ocorrência de um mesmo número de sucessos.
3. A probabilidade de obter dois ou mais sucessos em um intervalo suficientemente pequeno é desprezível.

Distribuição de Poisson

Seja t: comprimento total do intervalo

n: número de partes da divisão do intervalo, tal que no máximo um sucesso em cada parte

t/n: comprimento de cada parte do intervalo

Portanto:

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

Onde k: número de sucessos em n repartições

p: probabilidade de sucesso em cada parte

Seja λ: taxa de ocorrência de sucessos

(Ex.: chegadas/ hora; defeitos /metro)

Então:

$$\boxed{p = \frac{\lambda t}{n}} \Rightarrow \boxed{\Pr(X = k) = \binom{n}{k} \cdot \left(\frac{\lambda t}{n}\right)^k \cdot \left(1 - \frac{\lambda t}{n}\right)^{n-k}}$$

Considerando n ⟶ infinito : **(POISSON)**

$$\Pr(X = k) = \lim_{n \to \infty} \binom{n}{k} \cdot \left(\frac{\lambda t}{n}\right)^k \cdot \left(1 - \frac{\lambda t}{n}\right)^{n-k}$$

$$\boxed{\Pr(X = k) = \frac{e^{-\lambda t} \cdot (\lambda t)^k}{k!}}$$

Onde e = 2,71828... (Número de Euler)

Distribuição de Poisson

$$\boxed{E(X) = \sum_{k=0}^{\infty} k \frac{e^{-\lambda t}(\lambda t)^k}{k!} = \lambda t}$$

$$\boxed{Var(X) = \sum_{k=0}^{\infty} (k - \lambda t)^2 \frac{e^{-\lambda t}(\lambda t)^k}{k!} = \lambda t}$$

Exemplo: Num processo de fabricação de alumínio aparecem em média uma falha a cada 400 m (taxa de falha: λ= 0,0025 falhas/m).
Qual a probabilidade de ocorrer 3 falhas em 1000m?

E(x) = λt = 0,0025 falhas/m 1000 m = 2,5 falhas

k = 3

$$\Pr(X = k) = \frac{e^{-\lambda t} \cdot (\lambda t)^k}{k!}$$

$$\boxed{\Pr(X = 3) = \frac{e^{-2,5}\, 2,5^3}{3!} = 0,2138}$$

Distribuição Geométrica

Repetição de um experimento com distribuição de Bernoulli (sucesso ou fracasso) até obtenção do **primeiro sucesso**.

Condições do experimento:
- repetições independentes
- mesma probabilidade de sucesso *p*

$$\boxed{\Pr(X = k) = p \cdot q^{k-1},\ k = 1,2,3\ldots}$$

$$E(X) = \sum_{i} x_i \cdot \Pr(X = x_i) = \sum_{k=1}^{\infty} k \cdot p \cdot q^{k-1} = \frac{1}{p}$$

$$Var(X) = \sum_{i} \left[x_i - E(X)\right]^2 \cdot \Pr(X = x_i) = \sum_{k=1}^{\infty} \left(k - \frac{1}{p}\right)^2 \cdot p \cdot q^{k-1} = \frac{q}{p^2}$$

Distribuição de Pascal

Repetição de um experimento com distribuição de Bernoulli (sucesso ou fracasso) até obtenção do **r-ésimo sucesso.**

<u>Condições do experimento</u>:

- provas independentes
- mesma probabilidade de sucesso *p*

***r*-ésimo sucesso** ocorre na ***k*-ésima tentativa**

***k-1* tentativas anteriores** houve ***r-1* sucessos**

Daí

$$\Pr(X=k)=p\cdot\binom{k-1}{r-1}\cdot p^{r-1}\cdot q^{(k-1)-(r-1)}$$

$$\boxed{\Pr(X=k)=\binom{k-1}{r-1}\cdot p^{r}\cdot q^{k-r}}$$

Para: $k=r,\ r+1,\ r+2,\ \ldots$

Distribuição de Pascal

$$E(X) = \sum_{i} x_i \Pr(x_i)$$

$$E(X) = \sum_{k=r}^{\infty} k.\binom{k-1}{r-1} p^r q^{k-r} = \ldots =$$

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

$$Var(X) = \sum_{i} [x_i - E(X)]^2 \Pr(x_i)$$

$$Var(X) = \sum_{k=r}^{\infty} \left(k - \frac{r}{p}\right)^2 . \binom{k-1}{r-1} . p^r q^{k-r} = \ldots =$$

$$\boxed{Var(X) = \frac{r \cdot q}{p^2}}$$

Distribuição Hipergeométrica

Difere da Distribuição Binomial somente porque as repetições do experimento são feitas sem reposição.

Seja:

N: conjunto de elementos

r : subconjunto com determinada característica

n: elementos são extraídos sem reposição

X: número de elementos com tal característica

$$\Pr(X = k) = \frac{\binom{r}{k} \cdot \binom{N-r}{n-k}}{\binom{N}{n}}$$

$$E(X) = \sum_{k} k \cdot \frac{\binom{r}{k} \cdot \binom{N-r}{n-k}}{\binom{N}{n}} = \ldots = \frac{r \cdot n}{N} = n \cdot p$$

$$Var(X) = \sum_{k} (k - np)^2 \cdot \frac{\binom{r}{k} \cdot \binom{N-r}{n-k}}{\binom{N}{n}} = \ldots = n \cdot p \cdot q \cdot \frac{N-n}{N-1}$$

Distribuição Polinomial ou Multinomial

Condições do experimento:

- n repetições independentes
- cada repetição admite um único resultado dentre r possíveis resultados
- probabilidade de ocorrer um determinado resultado é constante
- X_i: número de ocorrências do i-ésimo resultado
- p_i: probabilidade de ocorrência do i-ésimo resultado

$$\boxed{\Pr(X_1 = k_1; X_2 = k_2; \ldots X_r = k_r) = \frac{n!}{k_1!\, k_2! \ldots k_r!} \cdot p_1^{k_1} \cdots p_r^{k_r}}$$

Onde:

$$\sum_{i=1}^{r} k_i = n \qquad \sum_{i=1}^{r} p_i = 1$$