1 Distribuições Discretas de Probabilidade

A distribuição discreta 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.

Exemplos

1. Lança-se uma moeda 10 vezes e anota-se o número de caras. Este número pode ser 0, 1, 2 ...10.
2. Em uma pesquisa de mercado feita com 200 pessoas, perguntam-se estes compram um determinado produto. O número de pessoas que compram o produto varia de 0 a 200.
3. Conta-se o número de acidentes que ocorrem em uma rodovia num feriado prolongado. O número de acidentes em questão pode ser: 0, 1, 2... Como não temos um valor que limite esse número, supomos que o número de acidentes é qualquer inteiro não negativo.
4. Número de chamadas telefônicas que chegam a uma central em um intervalo de tempo.

Existem várias distribuições discretas ou modelos probabilísticos discretos que podem ser usados em diversas situações práticas. O problema é determinar qual modelo é mais adequado para a situação em estudo, e como aplicá-lo adequadamente.

1.1 Distribuição Uniforme Discreta

É a mais simples das distribuições discretas e recebe o nome de uniforme porque todos os valores da variável aleatória são assumidos com a mesma probabilidade.

Exemplo o lançamento de um dado não viciado, definindo como X, a variável aleatória que representa a face voltada para cima, X assume os valores $x = 1, 2, 3, 4, 5, 6$ com a mesma probabilidade $1/6$.

A distribuição uniforme neste caso é dada por

$$f(x) = \frac{1}{6} \text{ para } x = 1, 2, 3, 4, 5, 6$$

Generalizado obtém-se a função de probabilidade

$$f(x) = \frac{1}{k} \text{ para } x = x_1, x_2, x_3, ..., x_k$$

k numero de termos.

Verifica-se então que $f(x)$ depende de k.

1.1.1 Parâmetros Característicos da Distribuição Uniforme

1. Média $\mu = \frac{k+1}{2}$
   No exemplo dos dados $\mu = \frac{6+1}{2} = 3,5$
2. Variância $\sigma^2 = \frac{k^2-1}{12}$
   No exemplo dos alérgicos $sigma^2 == \frac{6^2-1}{12} = 2,92$ s

1.2 Distribuição Bernoulli

Na prática existem muitos experimentos que admitem apenas dois resultados. Exemplos:

1. Uma peça é classificada como boa ou defeituosa;
2. Um entrevistado concorda ou não com a afirmação feita;
3. Um servidor de internet está ativo ou não;
4. Numa linha de produção observa-se se um item é defeituoso ou não.

Situações com alternativas dicotômicas podem ser representadas genericamente por respostas do tipo sucesso-fracasso.

Esses experimentos recebem o nome de ensaio de Bernoulli e originam uma variável aleatória com distribuição Bernoulli. Neste caso, consideramos uma experiência com dois possíveis resultados

- Sucesso $\rightarrow P(sucesso) = p$;
- Fracasso $\rightarrow P(fracasso) = q$.

Temos que:

$$\begin{aligned} \Omega &= \{Sucesso, Fracasso\} \therefore P(\Omega) = 1 \\ & p + q = 1 \quad q = 1 - p \end{aligned}$$

1.2.1 Parâmetros Característicos da Distribuição Uniforme

1. Média $\mu = p$
   No exemplo dos dados $\mu = \frac{6+1}{2} = 3,5$
2. Variância $\sigma^2 = pq$
   No exemplo dos alérgicos $sigma^2 == \frac{6^2-1}{12} = 2,92$ s

1.3 Distribuição Binomial

Na maior parte das vezes, são realizados n ensaios de Bernoulli. O interesse está no número X de ocorrências de sucessos.

Exemplos:

1. lançar uma moeda cinco vezes e observar o número de caras;
2. numa linha de produção, observar dez itens, e verificar quantos são defeituosos;
3. verificar, num dado instante, o número de processadores ativos, num sistema com multiprocessadores;

Uma experimento binomial é dado da seguinte forma:

1. consiste em n ensaios de Bernoulli;

2. cujos ensaios são independentes; e

3. para o qual a probabilidade de sucesso em cada ensaio é sempre igual a p, $0 < p < 1$

A variável aleatória X, correspondente ao número de sucessos num experimento binomial, tem distribuição binomial com parâmetros n e p, com função de probabilidade dada por

$$f(x) = C_x^n p^x q^{n-x}$$

A fórmula de cálculo de uma combinação é a seguinte:

$$C_x^n = \binom{n}{x} = \frac{n!}{x!\,(n-x)!}$$

A função $f(x)$ permite calcular a probabilidade de acontecer o resultado x (número de sucessos da variável aleatória), não importando a ordem de ocorrência de x dentro da experiência.

Exemplo: Numa família com $n = 5$ filhos, qual a probabilidade de não haver homens? Qual a probabilidade de haver dois homens? $n = 5, p = \frac{1}{2}, q = \frac{1}{2}$

$$f(x) = C_x^5 p^x q^{5-x}; \quad x = 0, 1, 2, 4, 5$$

A variável aleatória representa o número de homens (filhos do sexo masculino) encontrado em famílias de 5 filhos

1. $x = 0$ homem

$$\begin{aligned} f(x) &= C_0^5 p^0 q^{5-0} \\ &= \frac{5!}{0!\,(5-0)!}\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^5 \\ &= \frac{1}{32} = 0,0313 \,\text{ou}\, 3,13\% \end{aligned}$$

2. $x = 2$ homens

$$\begin{aligned} f(x) &= C_2^5 p^2 q^{5-2} \\ &= \frac{5!}{2!\,(5-2)!}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^3 \\ &= \frac{20}{2}\frac{1}{4}\frac{1}{8} = \frac{10}{32} = 0,3125 \,\text{ou}\, 31,25\% \end{aligned}$$

Exemplo: Lançada oito moedas (ou uma moeda oito vezes), qual a chance de obter

- Três caras?
- no máximo três caras?
- no mínimo quatro caras?

A variável aleatória x neste caso é o número de caras obtidos no lançamento, logo neste caso o sucesso sair cara nas moedas lançadas. Assim temos:

$$n = 8,\, p = \frac{1}{2} = 0,5\, q = 1 - q = 1 - 0,5 = 0,5$$

A função de probabilidade

$$f(x) = C_x^n p^x q^{n-x}$$

Probabilidade de sair três caras

$$\begin{aligned} P[X=3] &= C_3^8 p^3 q^{8-3} \\ &= \frac{8!}{3!\,(8-3)!}(0,5)^3(0,5)^5 \\ &= 56 \times 0,125 \times 0,03125 = 0,2187 \,\text{ou}\, 21,87\% \end{aligned}$$

Probabilidade de sair no máximo três caras

$$\begin{aligned} P[X \leq 3] &= P[X=0] + P[X=1] + P[X=2] + P[X=3] \\ P[X=0] &= C_0^8 p^0 q^{8-0} = 0,0039 \\ P[X=1] &= C_1^8 p^1 q^{8-1} = 0,0313 \\ P[X=2] &= C_2^8 p^2 q^{8-2} = 0,1094 \\ P[X=3] &= 0,2187 \\ P[X \leq 3] &= 0,0039 + 0,0313 + 0,1094 + 0,2187 = 0,3633 \,\text{ou}\, 36,33\% \end{aligned}$$

Probabilidade de sair no mínimo quatro caras

$$\begin{aligned} P[X \geq 4] &= P[X=4] + P[X=5] + P[X=6] + P[X=7] + P[X=8] \\ & \quad\text{ou} \\ P[X \geq 4] &= 1 - P[X < 4] = 1 - (P[X=0] + P[X=1] + P[X=2] + P[X=3]) \\ &= 1 - 0,3633 = 0,6367 \,\text{ou}\, 63,67\% \end{aligned}$$

1.3.1 Parâmetros Característicos da Distribuição Binomial

1. Média $\mu = np$

2. Variância $\sigma^2 = npq$

3. Desvio Padrão $\sigma = \sqrt{npq}$

Utilizando o exemplo das moedas temos:

1. Média $\mu = np = 8 \times 0,5 = 4$

2. Variância $\sigma^2 = 8 \times 0,5 \times 0,5 = 2$

3. Desvio Padrão $\sigma = \sqrt{npq} = \sqrt{2} = 1,41$

1.4 Distribuição Hipergeométrica

A distribuição hipergeométrica é intimamente relacionada à distribuição binomial. Enquanto a distribuição binomial é o modelo aproximado de amostragem sem reposição de uma população, dicotômica finita, a distribuição hipergeométrica é o modelo de probabilidade para o número de sucessos em uma amostra. As hipóteses que levam à distribuição hipergeométrica são as seguintes:

1. 1. A população ou o conjunto de onde é retirada a amostra consiste de N indivíduos, objetos ou elementos (população finita).

2. Cada indivíduo é classificado como sucesso (p) ou fracassos (q) e há M sucessos na população.

3. É selecionada uma amostra sem reposição de n indivíduos de forma que cada subconjunto de tamanho n seja igualmente provável de ser escolhido.

A distribuição hipergeométrica tem a seguinte função de probabilidade

$$f(x) = \frac{C_x^k C_{(n-x)}^{(N-k)},}{C_n^N} x = 0, 1, 2, 3, ...$$

em que:

- x é uma variável aleatória discreta;
- N quantidade de itens;
- n tamanho da amostra;
- k numero de sucessos;

Exemplo: Pequenos motores elétricos são expedidos em lotes de 50 unidades. Antes que uma remessa seja aprovada, um inspetor escolhe 5 desses motores e os inspeciona. Se nenhum dos motores inspecionados for defeituoso, o lote é aprovado. Se um ou mais forem verificados defeituosos, todos os motores da remessa são inspecionados. Suponha que existam, de fato, três motores defeituosos no lote. Qual a probabilidade de que a inspeção de todo o lote seja necessária?

Se fizermos igual a X o numero de motores defeituosos encontrados, inspeção de todo o lote seja necessária se $X \geq 1$

Neste caso temos $k = 3$ $n = 5$ $N = 50$;

$$\begin{aligned}
P[X = x] &= \frac{C_x^k C_{(n-x)}^{(N-k)},}{C_n^N} \\
P[X \geq 1] &= 1 - P[X < 1] = 1 - P[X = 0]) \\
P[X = 0] &= = \frac{C_0^3 C_{(5-0)}^{(50-3)},}{C_5^{50}} = \frac{C_0^3 C_5^{47},}{C_5^{50}} = 0,7239 \\
P[X \geq 1] &= 1 - 0,7239 = 0,2761
\end{aligned}$$

Quando se tem $\frac{n}{N} < 0,1$, pode-se utilizar a distribuição binomial para aproximar a distribuição hipergeométrica.

1.4.1 Parâmetros Característicos da Distribuição Hipergeométrica

1. Considerando $p = \dfrac{k}{N}$ e $q = 1 - p$
2. Média $\mu = np$
3. Variância $\sigma^2 = npq\dfrac{N-n}{N-1}$

1.5 Distribuição Geométrica

A distribuição geométrica está também associada à seqüência de uma prova de Bernoulli excetuando-se que o número de provas não é fixada, e, na verdade, a variável aleatória de interesse X é definida como o número de provas necessárias para obter o primeiro sucesso.

Exemplos:

- numero de vezes que uma pessoa estaciona num certo local proibido até apanhar uma multa;
- numero de tentativas até acertar no alvo (jogo de tiro ao alvo);
- numero de lançamentos de uma moeda até sair cara;

A distribuição geométrica tem a seguinte função de probabilidade

$$f(x) = pq^x \; x = 0, 1, 2, 3, ...$$

em que:

- x é uma variável aleatória discreta;
- p probabilidade de sucesso;
- q probabilidade de fracasso.

Exemplo: Se $0,05$ é a probabilidade de uma fábrica produzir uma peça defeituosa, qual é a probabilidade de pelo menos 2 peças boas sejam produzidas antes de se produzir a primeira defeituosa.

X o numero peças boas, então pelo menos 2 peças boas $X \geq 2$

Neste caso temos $p = 0,05$ $q = 0,95$;

$$\begin{aligned}
P[X = x] &= pq^x \\
P[X \geq 2] &= 1 - P[X < 2] = 1 - (P[X = 0] + P[X = 1]) \\
P[X = 0] &= (0,05)(0,95)^0 = 0,05 \\
P[X = 1] &= (0,05)(0,95)^1 = 0,0475 \\
P[X \geq 2] &= 1 - (0,05 + 0,0475) = 1 - 0,0975
\end{aligned}$$

1.5.1 Parâmetros Característicos da Distribuição Geométrica

1. Média $\mu = \dfrac{q}{p}$
2. Variância $\sigma^2 = \dfrac{q}{p^2}$

1.6 Distribuição de Poisson

A distribuição de Poisson é empregada em experimentos nos quais não se está interessado no número de sucessos obtido 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, comprimento, área, ou volume. Alguns exemplos de variáveis que podem ter a distribuição de Poisson são:

1. número de defeitos por centímetro quadrado;
2. número de acidentes por dia;
3. número de clientes por hora;
4. número de chamadas telefônicas recebidas por minuto;
5. número de falhas de um computador num dia de operação;
6. número de relatórios de acidentes enviados a uma companhia de seguros numa semana.

A distribuição de Poisson tem a seguinte função de probabilidade

$$f(x) = e^{-\lambda}\frac{\lambda^x}{x!},\ x = 0, 1, 2, 3, ...$$

em que:

- x é uma variável aleatória discreta;
- e base dos logaritmos neperianos (2,718...)
- λ - média da distribuição (λp)

Exemplo: O número médio de dias por ano que ocorrem chuvas acima de $50mm.h^{-1}$ em uma determinada região é 1,5. Qual a probabilidade de haver mais de dois dias com chuvas acima dessa intensidade.

$$\begin{aligned}
P[X = x] &= e^{-\lambda}\frac{\lambda^x}{x!} \\
P[X > 2] &= 1 - P[X \leq 2] = 1 - (P[X = 0] + P[X = 1] + P[X = 2]) \\
P[X = 0] &= = e^{-1,5}\frac{1,5^0}{0!} = 0,2231 \\
P[X = 1] &= = e^{-1,5}\frac{1,5^1}{1!} = 0,3347 \\
P[X = 2] &= = e^{-1,5}\frac{1,5^2}{2!} = 0,2510 \\
P[X > 2] &= 1 - (0,2231 + 0,3347 + 2510) = 1 - 0,8088 = 0,1912\, ou\, 19,12\%
\end{aligned}$$

A distribuição de Poisson também é conhecida na prática com lei dos eventos raros. Evento raro pode ser considerado quando $n \geq 50$ e $p \leq 0,10$.Nestes casos podemos utilizar a distribuição de Poisson para probabilidades de situações que seriam utilizadas uma distribuição binomial.

Exemplo: A probabilidade de que um indivíduo apresente reação alérgica após a aplicação de um soro é de 0,002. Esse mesmo soro foi aplicado a um grupo de 1800 pessoas, qual a probabilidade de que duas pessoas apresentem reação alérgica? n=1800 p=0,002

$\lambda = 1800 \times 0,002 = 3,6$ alérgicos

$$\begin{aligned}
P[X = x] &= e^{-\lambda}\frac{\lambda^x}{x!} \\
P[X = 2] &= = e^{-3,6}\frac{3,6^2}{2!}0,1770\text{ou}\, 17,70\%
\end{aligned}$$

1.6.1 Parâmetros Característicos da Distribuição de Poisson

1. Média $\mu = \lambda$

   No exemplo dos alérgicos $\mu = 3,6$

2. Variância $\sigma^2 = \lambda$

   No exemplo dos alérgicos $\sigma^2 = 3,6$ s

3. Desvio Padrão $\sigma = \sqrt{\lambda}$

   No exemplo da sementes $\sigma = \sqrt{3,6} = 1,9$