Iniciação à Estatística

Prof. Fernando de Souza Bastos
fernando.bastos@ufv.br

Departamento de Estatística
Universidade Federal de Viçosa
Campus UFV - Viçosa

Sumário

Modelos Probabilísticos

Algumas variáveis aleatórias adaptam-se muito bem a uma série de problemas práticos. Portanto, um estudo pormenorizado dessas variáveis é de grande importância para a construção de modelos probabilísticos para situações reais e a conseqüente estimação de seus parâmetros. Para algumas dessas distribuições existem tabelas que facilitam o cálculo de probabilidades, em função de seus parâmetros. Nesta seção iremos estudar alguns desses modelos, procurando enfatizar as condições em que eles aparecem, suas funções de probabilidade, parâmetros e como calcular probabilidades.

Distribuição Uniforme Discreta

Este é o caso mais simples de v.a. discreta, em que cada valor possível ocorre com a mesma probabilidade.

Definição: A v.a. discreta X, assumindo os valores $x_1, \cdots, x_k$, tem distribuição uniforme se, e somente se,

$$P(X = x_i) = p(x_i) = \frac{1}{k}, \quad \forall \quad i = 1, 2, \cdots, k.$$

É fácil verificar que,

$$E(X) = \frac{1}{k}\sum_{i=1}^{k} x_i, \quad Var(x) = \frac{1}{k}\left\{\sum_{i=1}^{k} x_i^2 - \frac{\left(\sum_{i=1}^{k} x_i\right)^2}{k}\right\}$$

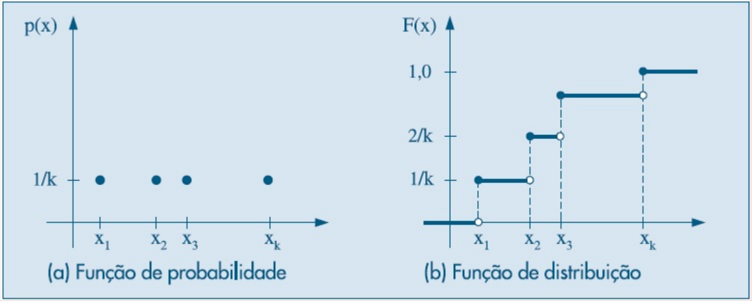

Figura: Distribuição uniforme discreta.

Exemplo

Seja X a v.a. que indica o "número de pontos marcados na face superior de um dado", quando ele é lançado. Obtemos na Tabela abaixo a distribuição de X.

x	1	2	3	4	5	6	Total
$p(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	1

Notemos que

$$E(X) = 3,5 \quad \text{e} \quad V(X) = 2,9$$

Distribuição Bernoulli

Muitos experimentos são tais que os resultados apresentam ou não uma determinada característica. Por exemplo:

1. uma moeda é lançada: o resultado ou é cara, ou não (ocorrendo, então, coroa);

Distribuição Bernoulli

Muitos experimentos são tais que os resultados apresentam ou não uma determinada característica. Por exemplo:

1. uma moeda é lançada: o resultado ou é cara, ou não (ocorrendo, então, coroa);
2. um dado é lançado: ou ocorre face 5 ou não (ocorrendo, então, uma das faces 1, 2, 3, 4 ou 6);

Distribuição Bernoulli

Muitos experimentos são tais que os resultados apresentam ou não uma determinada característica. Por exemplo:

1. uma moeda é lançada: o resultado ou é cara, ou não (ocorrendo, então, coroa);
2. um dado é lançado: ou ocorre face 5 ou não (ocorrendo, então, uma das faces 1, 2, 3, 4 ou 6);
3. uma peça é escolhida ao acaso de um lote contendo 500 peças: essa peça é defeituosa ou não;

Distribuição Bernoulli

Muitos experimentos são tais que os resultados apresentam ou não uma determinada característica. Por exemplo:

1. uma moeda é lançada: o resultado ou é cara, ou não (ocorrendo, então, coroa);
2. um dado é lançado: ou ocorre face 5 ou não (ocorrendo, então, uma das faces 1, 2, 3, 4 ou 6);
3. uma peça é escolhida ao acaso de um lote contendo 500 peças: essa peça é defeituosa ou não;
4. uma pessoa escolhida ao acaso dentre 1.000 é ou não do sexo masculino;

Distribuição Bernoulli

Muitos experimentos são tais que os resultados apresentam ou não uma determinada característica. Por exemplo:

1. uma moeda é lançada: o resultado ou é cara, ou não (ocorrendo, então, coroa);
2. um dado é lançado: ou ocorre face 5 ou não (ocorrendo, então, uma das faces 1, 2, 3, 4 ou 6);
3. uma peça é escolhida ao acaso de um lote contendo 500 peças: essa peça é defeituosa ou não;
4. uma pessoa escolhida ao acaso dentre 1.000 é ou não do sexo masculino;
5. uma pessoa é escolhida ao acaso entre os moradores de uma cidade e verifica-se se ela é favorável ou não a um projeto municipal.

Em todos esses casos, estamos interessados na ocorrência de sucesso (cara, face 5 etc.) ou fracasso (coroa, face diferente de 5 etc.). Essa terminologia (sucesso e fracasso) será usada freqüentemente.

Em todos esses casos, estamos interessados na ocorrência de sucesso (cara, face 5 etc.) ou fracasso (coroa, face diferente de 5 etc.). Essa terminologia (sucesso e fracasso) será usada freqüentemente.

Para cada experimento acima, podemos definir uma v.a. X, que assume apenas dois valores: 1, se ocorrer sucesso, e 0, se ocorrer fracasso. Indicaremos por p a probabilidade de sucesso, isto é, $P(sucesso) = P(S) = p, 0 < p < 1$.

Definição: A variável aleatória X, que assume apenas os valores 0 e 1, com função de probabilidade $(x, p(x))$ tal que

$$
\begin{aligned}
p(0) &= P(X = 0) = 1 - p,\\
p(1) &= P(X = 1) = p,
\end{aligned}
$$

é chamada variável aleatória de Bernoulli.

Definição: A variável aleatória X, que assume apenas os valores 0 e 1, com função de probabilidade $(x, p(x))$ tal que

$$p(0) = P(X = 0) = 1 - p,$$
$$p(1) = P(X = 1) = p,$$

é chamada variável aleatória de Bernoulli.

Observação: Experimentos que resultam numa v.a. de Bernoulli são chamados ensaios de Bernoulli. Usaremos a notação

$$X \sim Ber(p)$$

para indicar uma v.a. com distribuição de Bernoulli com parâmetro p.

Segue-se facilmente que

$$E(X) = p \quad Var(X) = p - p^2 = p(1-p)$$

e,

$$F(x) = \begin{cases} 0, & \text{se} \quad x < 0; \\ 1-p, & \text{se} \quad 0 \leq x < 1; \\ 1, & \text{se} \quad x \geq 1. \end{cases}$$

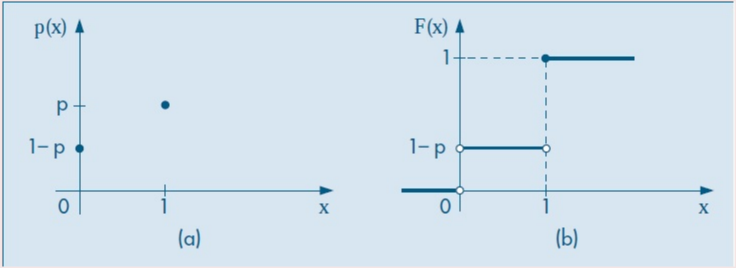

Figura: Distribuição de Bernoulli (a) f.p. (b) f.d.a.

Distribuição Binomial

Imagine, agora, que repetimos um ensaio de Bernoulli n vezes, ou, de maneira alternativa, obtemos uma amostra de tamanho n de uma distribuição de Bernoulli. Suponha ainda que as repetições sejam independentes, isto é, o resultado de um ensaio não tem influência nenhuma no resultado de qualquer outro ensaio. Uma amostra particular será constituída de uma seqüência de sucessos e fracassos, ou, alternativamente, de uns e zeros.

Distribuição Binomial

Por exemplo, repetindo um ensaio de Bernoulli cinco vezes ($n = 5$), um particular resultado pode ser FSSFS ou a quíntupla ordenada $(0, 1, 1, 0, 1)$. Usando a notação $P(S) = p$, a probabilidade de tal amostra será:

$$(1 - p)pp(1 - p)p = p^3 * (1 - p^2)$$

O número de sucessos nessa amostra é igual a 3, sendo 2 o número de fracassos.

Distribuição Binomial

Designamos por X o número total de sucessos em n ensaios de Bernoulli, com probabilidade de sucesso $p, 0 < p < 1$. Os possíveis valores de X são $0, 1, 2, ..., n$ e os pares $(x, p(x))$, onde $p(x) = P(X = x)$, constituem a chamada distribuição binomial.

Distribuição Binomial

Assim, numa seqüência de n ensaios de Bernoulli, a probabilidade de obter x sucessos (e portanto $n-x$ fracassos), $x = 0, 1, 2, ..., n$, com $P(S) = p, P(F) = 1-p = q$, é dado por $p^x(1-p)^{n-x} = p^x q^{n-x}$, devido à independência dos ensaios. Mas qualquer seqüência com x sucessos e $n-x$ fracassos terá a mesma probabilidade. Portanto resta saber quantas seqüências com a propriedade especificada podemos formar. É fácil ver que existem

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

logo,

$$P(X = x) = \binom{n}{x} p^x q^{n-x}, \; x = 0, 1, 2, ..., n.$$

Distribuição Binomial

Se X tem distribuição binomial com parâmetros n e p, indicamos $X \sim Bin(n, p)$. Nesse caso,

$$E(X) = np;$$
$$V(X) = np(1-p)$$

Distribuição de Poisson

Dizemos que uma v.a. N tem uma distribuição de Poisson com parâmetro $\lambda > 0$ se,

$$P(N = k) = \frac{\exp(-\lambda)\lambda^k}{k!}, \ k = 0, 1, 2, \cdots,$$

Neste caso, $E(N) = Var(N) = \lambda$; Logo, λ representa o número médio de eventos ocorrendo no intervalo considerado.

A distribuição de Poisson é largamente empregada quando se deseja contar o número de eventos de certo tipo que ocorrem num intervalo de tempo, ou superfície ou volume. São exemplos:

- número de chamadas recebidas por um telefone durante cinco minutos;
- número de falhas de um computador num dia de operação; e
- número de relatórios de acidentes enviados a uma companhia de seguros numa semana.

Exemplo

Um telefone recebe, em média, cinco chamadas por minuto. Supondo que a distribuição de Poisson seja adequada nessa situação, obter a probabilidade de que o telefone não receba chamadas durante um intervalo de um minuto.

Exemplo

Um telefone recebe, em média, cinco chamadas por minuto. Supondo que a distribuição de Poisson seja adequada nessa situação, obter a probabilidade de que o telefone não receba chamadas durante um intervalo de um minuto.

Resolução:

$N =$"Número de chamadas por minuto"

Notemos que $\lambda = E(N) = 5$, portanto:

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

Por outro lado, se quisermos a probabilidade de obter no máximo duas chamadas em quatro minutos, teremos $\lambda = 20$ chamadas em quatro minutos, logo

$$
\begin{aligned}
P(N \leq 2) &= P(N=0) + P(N=1) + P(N=2) \\
&= \exp(-20)(1 + 20 + 200) \\
&= 221 \exp(-20)
\end{aligned}
$$

Denotaremos uma v.a. N com distribuição de Poisson de parâmetro λ por:

$$N \sim Poisson(\lambda)$$

Referências