WIKIPÉDIA

Função distribuição acumulada

Origem: Wikipédia, a enciclopédia livre.

Em teoria da probabilidade, a **função distribuição acumulada** (**fda**) ou simplesmente **função distribuição**, descreve completamente a distribuição da probabilidade de uma variável aleatória de valor real *X*. Para cada número real *x*, a fda é dada por[1]:

Em linguagem matemática	Em Português
$F(x) = \mathrm{P}(X \leq x),$	A função de nome "F" é igual à probabilidade de que a variável aleatória *X* assuma um valor inferior ou igual a determinado *x*. Note que, via de regra, para cada *x*, a função F assumirá um valor diferente.

A probabilidade de que *X* se situe num intervalo $]a, b]$ (aberto em *a* e fechado em *b*) é $F(b) - F(a)$ se $a \leq b$. É convenção usar um *F* maiúsculo para a fda, em contraste com o *f* minúsculo usado para a função densidade da probabilidade e função massa de probabilidade.

A função distribuição pode ser facilmente obtida a partir da função de probabilidade respectiva. No caso duma variável aleatória discreta:

$$F(x) = \sum_{x_i \leq x} f(x_i)$$

Para uma variável aleatória contínua:

$$F(x) = \int_{-\infty}^{x} f(x_i)\, dx$$

Note-se que na definição acima, o sinal "menor ou igual", '≤' poderia ser substituído por "menor" '<'. Isto produziria uma função diferente, mas qualquer uma das funções pode ser facilmente deduzida a partir da outra. Também se poderia mudar para um sinal maior e deduzir as propriedades desta nova função. A única coisa a lembrar é ajustar a definição ao sinal pretendido. Em países de língua inglesa, a convenção que usa a desigualdade fraca (≤) em vez da desigualdade estrita (<) é quase sempre usada.

Índice

Exemplos

Como exemplo, suponha-se que *X* é distribuído uniformemente pelo intervalo [0, 1]. Nesse caso a fda é dada por:

$F(x) = 0$, se $x < 0$;
$F(x) = x$, se $0 \leq x \leq 1$;
$F(x) = 1$, se $x > 1$.

Para um outro exemplo suponha-se que *X* toma apenas os valores 0 e 1, com igual probabilidade (*X* segue a distribuição de Bernoulli com p = 1/2). Então a fda é dada por

$F(x) = 0$, se $x < 0$;
$F(x) = 1/2$, se $0 \leq x < 1$;
$F(x) = 1$, se $x \geq 1$.

Notação

Quando há mais de uma variável aleatória e torna-se necessário explicitar a diferença entre as funções, representa-se a fda da variável aleatória *X* por $\mathbf{F}_X(x)$.

Propriedades

Se *X* é uma variável aleatória discreta, então ela obtém os valores x_1, x_2, ... com probabilidade p_1, p_2 etc., e a fda de *X* será descontínua nos pontos x_i e constante entre eles.

Se a fda *F* de *X* é contínua, então *X* é uma variável aleatória contínua; se para além disso *F* absolutamente contínua, então existe uma função Integral Lebesgue *f*(*x*) tal que

$$F(b) - F(a) = \mathrm{P}(a \leq X \leq b) = \int_a^b f(x)\,dx$$

para todos os números reais *a* e *b*. A primeira das duas igualdades acima não seria correcta em geral se não tivéssemos dito que a distribuição é contínua. Continuidade da distribuição implica que P(*X* = *a*) = P(*X* = *b*) = 0, de modo que a diferença entre "<" e "≤" deixa de ser importante neste contexto.) A função *f* é igual à derivada de *F* (quase em toda a parte), e é chamada de função densidade de probabilidade da distribuição de *X*.

Para qualquer função de distribuição F, tem-se:

- $0 \leq F(x) \leq 1$
- F é não decrescente (crescente ou constante): $x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2)$
- $F(-\infty) = \lim\limits_{x \to -\infty} F(x) = 0$
- $F(+\infty) = \lim\limits_{x \to +\infty} F(x) = 1$
- F é contínua à direita: $F(a^+) = \lim\limits_{x \to a^+} F(x) = F(a)$
- $\mathrm{P}(x = a) = F(a) - F(a^-)$
- $\mathrm{P}(a < x \leq b) = F(b) - F(a)$, com $a, b \in \mathbb{R}$, e $a < b$

Temos ainda as seguintes propriedades, que permitem lidar com os diferentes tipos de desigualdades, e que se aplicam a funções distribuição de variáveis aleatórias discretas:

- $\mathrm{P}(X < b) = F(b^-)$
- $\mathrm{P}(X > a) = 1 - F(a)$
- $\mathrm{P}(X \geq a) = 1 - F(a^-)$
- $\mathrm{P}(a < X < b) = F(b^-) - F(a)$
- $\mathrm{P}(a \leq X < b) = F(b^-) - F(a^-)$
- $\mathrm{P}(a \leq X \leq b) = F(b) - F(a^-)$

No caso das variáveis aleatórias contínuas, valem as seguintes propriedades:

- F é contínua em todos os pontos (no caso das v. a. discretas era apenas contínua à direita)
- $\mathrm{P}(x = a) = \int_a^a f(x)\, dx = 0$
- $$\mathrm{P}(a \leq X \leq b) = \mathrm{P}(a \leq X < b) = \mathrm{P}(a < X \leq b) = \mathrm{P}(a < X < b) = \int_a^b f(x)\, dx = F(b) - F(a)$$

O teste Kolmogorov-Smirnov é baseado em funções distribuição acumulada e pode ser usado para ver se duas distribuições empíricas são diferentes ou se uma distribuição empírica é diferente de uma distribuição ideal. Muito relacionado é o teste de Kuiper, o qual é útil se o domínio da distribuição é cíclico como por exemplo em dias da semana. Por exemplo podemos usar o teste de Kuiper para ver se o número de tornados varia durante o ano ou se as vendas de um produto variam dia a dia ou por dia do mês.

Ver também

- Estatística descritiva
- Distribuição de probabilidade

Referências

1. Distribution and Quantile Functions (http://www.math.uah.edu/stat/dist/CDF.xhtml), *site* www.math.uah.edu, do Department of Mathematical Sciences da University of Alabama in Huntsville

Bibliografia

- Conceito de variável aleatória e de função de distribuição (http://w3.ualg.pt/~sjesus/aulas/pds/node108.html)
- Portal Action (http://www.portalaction.com.br/content/21-fun%C3%A7%C3%A3o-de-distribui%C3%A7%C3%A3o-acumulada)