

MAT 461 – Tópicos de Matemática II

Aula 7: Resumo de Probabilidade

Edson de Faria

Departamento de Matemática
IME-USP

28 de Agosto, 2013

1 Variáveis Aleatórias Contínuas

- Definição e Exemplos
- Distribuição acumulada de uma variável aleatória

Variáveis Aleatórias Contínuas

- Nesta aula introduziremos o importante conceito de **variável aleatória contínua**, e exploraremos suas propriedades.
- Tais variáveis aparecem naturalmente, não apenas na modelagem de fenômenos envolvendo grandezas aleatórias com um espectro contínuo de valores, mas também como **limites de médias de variáveis aleatórias discretas**.

Variáveis Aleatórias Contínuas: motivação

A título de motivação, consideremos o seguinte exemplo.

Exemplo (1)

Alice e Beto marcaram um encontro para amanhã à noite, entre 9:00 e 10:00 horas, em frente ao prédio da Gazeta. Sendo ambos pessoas a um só tempo impontuais e impacientes, combinaram que o primeiro a chegar esperará apenas 10 minutos; se o segundo não chegar nesse intervalo de tempo, o primeiro irá embora. Qual é a probabilidade de que o encontro se concretize?

- Este é um exemplo típico de situação que pode ser modelada com a introdução de variáveis aleatórias contínuas.
- De fato, se X e Y representam os instantes de chegada ao local do encontro de Alice e Beto, respectivamente, podemos considerar X e Y como variáveis aleatórias cujos valores se distribuem de maneira **uniforme** no intervalo de tempo que vai das 9:00 às 10:00 horas.
- A solução deste exemplo será apresentada mais adiante.

Variáveis Aleatórias Contínuas: definição

- Consideremos a definição formal de variável aleatória contínua.

Definição

Dizemos que uma variável aleatória $X : \Omega \rightarrow \mathbb{R}$ sobre um espaço de probabilidade $(\Omega, \mathcal{A}, P)$ é **contínua** se existe uma função $f_X : \mathbb{R} \rightarrow \mathbb{R}$ integrável não-negativa tal que, para quaisquer a, b reais com $a < b$, temos

$$P[a < X < b] = \int_a^b f_X(t) dt .$$

A função f_X é chamada de **densidade de probabilidade** da variável aleatória X .

Note que

$$\int_{-\infty}^{+\infty} f_X(t) dt = P[-\infty < X < +\infty] = 1 .$$

Exemplo (Distribuição uniforme)

Dizemos que uma variável aleatória X é *uniformemente distribuída* no intervalo $(0, 1)$ se X tem uma densidade f_X dada por

$$f_X(t) = \begin{cases} 1 & \text{se } 0 < t < 1 \\ 0 & \text{caso contrário} \end{cases}$$

- Note que, se X é uniformemente distribuída no intervalo $(0, 1)$, e $0 < a < b < 1$, temos

$$P[a < X < b] = \int_a^b f_X(t) dt = \int_a^b 1 \cdot dt = b - a$$

Exemplo (Distribuição uniforme)

Mais geralmente, dizemos que uma variável aleatória X é **uniformemente distribuída** no intervalo (α, β) se X tem uma densidade f_X dada por

$$f_X(t) = \begin{cases} \frac{1}{\beta - \alpha} & \text{se } \alpha < t < \beta \\ 0 & \text{caso contrário} \end{cases}$$

- Note que, se X é uniformemente distribuída no intervalo (α, β) , e $\alpha < a < b < \beta$, temos

$$P[a < X < b] = \int_a^b f_X(t) dt = \int_a^b \frac{1}{\beta - \alpha} \cdot dt = \frac{b - a}{\beta - \alpha}$$

Exemplo (Distribuição exponencial)

Dizemos que uma variável aleatória X é *exponencial com parâmetro $\lambda > 0$* se X tem uma densidade f_X dada por

$$f_X(t) = \begin{cases} \lambda e^{-\lambda t} & \text{se } t \geq 0 \\ 0 & \text{se } t < 0 \end{cases}$$

- A distribuição exponencial é comumente utilizada para modelar situações em que estamos interessados no **intervalo de tempo** até a ocorrência de algum evento, tipicamente raro.
- Por exemplo, o tempo a partir de agora até a ocorrência do próximo abalo sísmico. Ou o tempo a partir de agora até o recebimento da próxima chamada telefônica.
- Num certo sentido, podemos dizer que a distribuição exponencial é um caso limite da distribuição de Poisson.

Exemplo (2)

Suponha que o tempo de funcionamento de uma lâmpada (em horas) até queimar é uma variável aleatória exponencial X com densidade dada por

$$f_X(t) = \begin{cases} \lambda e^{-t/1000} & \text{se } t \geq 0 \\ 0 & \text{se } t < 0 \end{cases}$$

para um certo $\lambda > 0$. Qual é a probabilidade de que

- A lâmpada funcione entre 500 e 1500 horas até queimar?
- A lâmpada funcione por menos de 1000 até queimar?

Solução:

- Como devemos ter

$$1 = \int_{-\infty}^{\infty} f_X(t) dt = \lambda \int_0^{\infty} e^{-t/1000} dt ,$$

concluimos facilmente que $1000\lambda = 1$, ou seja, $\lambda = 1/1000$.

- Portanto, temos:

$$\begin{aligned} P[500 < X < 1\,500] &= \int_{500}^{1\,500} \frac{1}{1\,000} e^{-t/1\,000} dt \\ &= -e^{-t/1\,000} \Big|_{500}^{1\,500} = e^{-\frac{1}{2}} - e^{-\frac{3}{2}} \simeq 0.384 . \end{aligned}$$

Isto responde a primeira pergunta do enunciado.

- Analogamente, temos

$$\begin{aligned} P[X < 1\,000] &= \int_0^{1\,000} \frac{1}{1\,000} e^{-t/1\,000} dt = -e^{-t/1\,000} \Big|_0^{1\,000} \\ &= 1 - e^{-1} \simeq 0.633 . \end{aligned}$$

- Em outras palavras, há uma probabilidade de 63.3% de que a lâmpada funcione no máximo 1 000 horas até queimar, e isto responde a segunda pergunta do enunciado.

Exemplos de densidades

Não é exagero dizer que a distribuição definida a seguir é a mais importante em modelagem probabilística.

Exemplo (Distribuição normal)

Dizemos que uma variável aleatória X é *normal com parâmetros* $\mu \in \mathbb{R}$ e $\sigma > 0$ se X tem uma densidade f_X dada por

$$f_X(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$$

- Para que esta definição seja consistente, é necessário que

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt = 1.$$

- Isto de fato é verdade, e se segue da identidade:

$$\int_{-\infty}^{\infty} e^{-y^2/2} dy = \sqrt{2\pi}.$$

Exemplos de densidades

- A distribuição normal aparece naturalmente na modelagem dos mais variados fenômenos, sendo a mais útil das distribuições probabilísticas.
- O principal motivo é o fato de que as médias de uma sequência de variáveis aleatórias independentes e igualmente distribuídas convergem para uma distribuição normal.
- Este fato, cujo enunciado correto será dado mais tarde, é conhecido como o **Teorema Central do Limite**.
- Veremos adiante que os parâmetros μ e σ de uma distribuição normal são exatamente a média e o desvio padrão, respectivamente, da variável aleatória normal correspondente.
- Uma variável aleatória normal com $\mu = 0$ e $\sigma = 1$ é chamada de **variável normal padrão**.

Definição

A *distribuição acumulada* de uma variável aleatória X é a função $F_X : \mathbb{R} \rightarrow \mathbb{R}$ dada por

$$F_X(x) = P[X \leq x] .$$

- Observamos que, se X é uma variável aleatória **contínua** com densidade f_X , então:

$$F_X(x) = P[-\infty < X \leq x] = \int_{-\infty}^x f_X(t) dt$$

- Assim, vemos que a distribuição acumulada de uma variável aleatória contínua X é uma função derivável, com $F'_X(x) = f_X(x)$.

Distribuição acumulada de uma variável aleatória

- Quando conhecemos a distribuição acumulada de uma variável aleatória, podemos calcular facilmente certas probabilidades associadas àquela variável.
- Por exemplo, temos

$$P[a < X \leq b] = F_X(b) - F_X(a)$$

- Quais são as distribuições acumuladas das variáveis aleatórias introduzidas acima?
- Se X é uma variável aleatória uniforme no intervalo (α, β) , é fácil ver que

$$F_X(x) = \begin{cases} 0 & \text{se } x \leq \alpha \\ \frac{x - \alpha}{\beta - \alpha} & \text{se } \alpha < x < \beta \\ 1 & \text{se } x \geq \beta \end{cases}$$

- Por outro lado, se X é uma variável aleatória exponencial com parâmetro $\lambda > 0$, então

$$F_X(x) = \begin{cases} 0 & \text{se } x \leq 0 \\ 1 - e^{-\lambda x} & \text{se } x > 0 \end{cases}$$

- Quando X é uma variável aleatória normal, infelizmente não podemos calcular explicitamente a integral que nos fornece a distribuição acumulada, a saber

$$F_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt$$

Isto se deve ao fato de que a função e^{-y^2} não possui primitiva que possa ser expressa em termos de funções elementares. Mas é perfeitamente possível calcular, para qualquer x dado, o valor numérico aproximado da integral acima.

- Muitos valores da distribuição acumulada da normal **padrão** encontram-se tabelados na maioria dos bons livros de probabilidade e/ou estatística.¹

¹Veja por exemplo S. Ross, *A first course in Probability*, 8th. ed., Prentice Hall (2010), página 201.

Esperança de uma variável aleatória contínua

Por analogia com o caso discreto, definimos a esperança de uma variável aleatória contínua da seguinte maneira.

Definição

Se X é uma variável aleatória contínua com densidade f_X , definimos a esperança de X por

$$E[X] = \int_{-\infty}^{\infty} t f_X(t) dt ,$$

desde que, obviamente, a integral no segundo membro exista.

Exemplo

Seja X uma variável aleatória contínua cuja densidade é dada por:

$$f_X(t) = \begin{cases} 2t & \text{se } 0 \leq t \leq 1 \\ 0 & \text{caso contrário} \end{cases}$$

Calcule $E[X]$.

Solução: Temos

$$E[X] = \int_{-\infty}^{\infty} t f_X(t) dt = \int_0^1 2t^2 dt = \frac{2}{3}.$$

Às vezes, para calcularmos $E[X]$ precisamos primeiramente conhecer a distribuição acumulada de X . Vejamos um exemplo simples.

Exemplo

Seja X uma variável aleatória uniformemente distribuída no intervalo $[0, 1]$. Calcule $E[e^X]$.

Solução:

- Seja $Y = e^X$; queremos calcular $E[Y]$.
- Lembremos que a densidade de X é dada por

$$f_X(t) = \begin{cases} 1 & \text{se } 0 \leq t \leq 1 \\ 0 & \text{caso contrário} \end{cases}$$

- Calculemos primeiramente a distribuição acumulada de Y , a saber $F_Y(x) = P[Y \leq x]$.

- Observe que X nunca assume valores negativos; portanto, Y não assume valores < 1 , ou seja, $P[Y \leq x] = 0$, para todo $x < 1$.
- Quando $1 \leq x \leq e$, temos por outro lado

$$\begin{aligned} P[Y \leq x] &= P[e^X \leq x] = P[X \leq \ln x] \\ &= \int_0^{\ln x} f_X(t) dt = \int_0^{\ln x} 1 dt = \ln x. \end{aligned}$$

- Observe que $P[Y \leq e] = 1$; portanto, $P[Y \leq x] = 1$ para todo $x \geq e$.
- Assim, vemos que

$$F_Y(x) = \begin{cases} 0 & \text{se } x < 1 \\ \ln x & \text{se } 1 \leq x \leq e \\ 1 & \text{se } x > e \end{cases}$$

- Derivando a função distribuição acumulada, obtemos a densidade de Y ; ou seja, $f_Y(t) = F'_Y(t)$.
- Neste caso específico, como $(\ln x)' = \frac{1}{x}$, obtemos

$$f_Y(t) = \begin{cases} \frac{1}{t} & \text{se } 1 \leq t \leq e \\ 0 & \text{caso contrário} \end{cases}$$

- Portanto, concluímos finalmente que

$$\begin{aligned} E[e^X] &= E[Y] = \int_{-\infty}^{\infty} t f_Y(t) dt \\ &= \int_1^e t \cdot \frac{1}{t} dt = \int_1^e dt = e - 1 \simeq 1.718 . \end{aligned}$$

O cálculo efetuado no exemplo acima poderia ter sido obtido de maneira mais direta mediante a seguinte proposição geral.

Proposição

Seja X uma variável aleatória contínua com densidade f_X , e seja $g : \mathbb{R} \rightarrow \mathbb{R}$ uma função real. Então

$$E[g(X)] = \int_{-\infty}^{\infty} g(t) f_X(t) dt$$

No exemplo anterior, temos $g(t) = e^t$. Portanto, aplicando a proposição acima temos

$$E[e^X] = \int_{-\infty}^{\infty} e^t f_X(t) dt = \int_0^1 e^t dt = e - 1 ,$$

em absoluta concordância com o resultado obtido anteriormente.

Variância de uma variável aleatória contínua

- A **variância** de uma variável aleatória contínua X é definida de maneira inteiramente análoga ao caso discreto:

$$\text{Var}(X) = E[(X - \mu)^2],$$

onde $\mu = E[X]$ é a média (ou valor esperado) de X .

- Portanto, pela proposição acima, temos:

$$\text{Var}(X) = \int_{-\infty}^{\infty} (t - \mu)^2 f_X(t) dt.$$

- Como no caso discreto, vale a identidade $\text{Var}(X) = E[X^2] - (E[X])^2$.
- Assim, quando dispomos da média μ , podemos, para obter a variância, simplesmente calcular:

$$E[X^2] = \int_{-\infty}^{\infty} t^2 f_X(t) dt$$

e subtrair μ^2 do resultado obtido.

Analisemos os três principais exemplos de variáveis aleatórias contínuas vistos até agora.

- **Distribuição uniforme:** Seja X uniformemente distribuída no intervalo (α, β) .
 - A esperança de X é:

$$\begin{aligned}\mu = E[X] &= \int_{-\infty}^{\infty} t f_X(t) dt = \int_{\alpha}^{\beta} \frac{t}{\beta - \alpha} dt \\ &= \frac{1}{2} \frac{\beta^2 - \alpha^2}{\beta - \alpha} = \frac{\alpha + \beta}{2} .\end{aligned}$$

- A variância de X é obtida calculando-se:

$$\begin{aligned}E[X^2] &= \int_{-\infty}^{\infty} t^2 f_X(t) dt = \int_{\alpha}^{\beta} \frac{t^2}{\beta - \alpha} dt \\ &= \frac{1}{3} \frac{\beta^3 - \alpha^3}{\beta - \alpha} = \frac{\alpha^2 + \alpha\beta + \beta^2}{3} .\end{aligned}$$

- Portanto,

$$\text{Var}(X) = E[X^2] - \mu^2 = \frac{\alpha^2 + \alpha\beta + \beta^2}{3} - \left(\frac{\alpha + \beta}{2}\right)^2 = \frac{(\alpha - \beta)^2}{12}$$

- **Distribuição exponencial:** Seja X uma variável aleatória exponencial com parâmetro $\lambda > 0$.

- A esperança de X é:

$$\mu = E[X] = \int_{-\infty}^{\infty} t f_X(t) dt = \int_0^{\infty} \lambda t e^{-\lambda t} dt$$

Integrando por partes, obtemos:

$$\mu = E[X] = \int_0^{\infty} e^{-\lambda t} dt = \left(\frac{e^{-\lambda t}}{-\lambda} \right) \Big|_0^{\infty} = \frac{1}{\lambda}$$

- A variância de X é obtida calculando-se:

$$\begin{aligned} E[X^2] &= \int_{-\infty}^{\infty} t^2 f_X(t) dt = \int_0^{\infty} \lambda t^2 e^{-\lambda t} dt \\ &= \left(-t^2 e^{-\lambda t} \right) \Big|_0^{\infty} + 2 \int_0^{\infty} t e^{-\lambda t} dt = \frac{2\mu}{\lambda} = \frac{2}{\lambda^2} \end{aligned}$$

Ou seja, temos:

$$\text{Var}(X) = E[X^2] - \mu^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

- **Distribuição normal:** Seja agora X uma variável aleatória normal, com parâmetros $\mu \in \mathbb{R}$ e $\sigma > 0$. Mostraremos em aula que $E[X] = \mu$ e que $\text{Var}(X) = \sigma^2$.