Universidade de São Paulo
Instituto de Física de São Carlos
SFI 5704 - Mecânica Estatística - 2021-1

Prof. Leonardo Paulo Maia

Lista 03 - 2021/04/20 → **2021/04/28** (analíticos) e **2021/05/03** (computacionais)

Calcule os cumulantes das variáveis aleatórias discretas dos itens (a) e (b). Para cada variável aleatória contínua indicada nas questões de (c) até (e), calcule sua função característica e, a partir dela, calcule a média e a variância da v.a.. Nos casos (d) e (e), calcule também os cumulantes.

a. Bernoulli, $p_0 = 1 - p$, $p_1 = p$

b. binomial, $p_n = \binom{N}{n} p^n (1-p)^{N-n}$, $n = 0, \cdots, N$

c. uniforme, $\rho(x) = 1/(b-a)$, $a \leq x \leq b$

d. exponencial, $\rho(x) = \lambda \mathrm{e}^{-\lambda x}$, $x \geq 0$

e. gama, $\rho(x) = \lambda^p x^{p-1} \mathrm{e}^{-\lambda x} / \Gamma(p)$, $x \geq 0$, $\Gamma(p) = \int_0^\infty z^{p-1} \mathrm{e}^{-z}\, dz$

f. (peso 2) Considere sequências binárias de tamanho n onde, em cada posição, o dígito 1 ocorre com probabilidade p e o dígito 0 ocorre com probabilidade $1-p$, independentemente do que ocorre nas demais posições. Use a lei da probabilidade total para obter uma equação de recorrência de segunda ordem para a probabilidade de uma dessas sequências **não apresentar dois 1's consecutivos**. Você deverá obter a solução explícita da recorrência, considerando as condições iniciais pertinentes.

g. (peso 2) Considere sequências binárias de tamanho n onde, em cada posição, o dígito 1 ocorre com probabilidade p e o dígito 0 ocorre com probabilidade $1-p$, independentemente do que ocorre nas demais posições. Use a lei da probabilidade total para obter uma equação de recorrência de primeira ordem para a probabilidade de uma dessas sequências apresentar uma quantidade total **par de 1's**. Você deverá obter a solução explícita da recorrência, considerando as condições iniciais pertinentes. Porém, antes de mais nada, tente predizer o comportamento qualitativo da solução. Para obter um número par de 1's, você preferiria p alto, baixo ou seria indiferente? A paridade de n é relevante? Sempre?

h. Mostre que $\boxed{\mathrm{cov}(X,Y) = \langle X.Y \rangle - \langle X \rangle . \langle Y \rangle}$ e que $\mathrm{cov}(X,Y) = 0$ se X e Y forem independentes.

i. Mostre que $\boxed{\mathrm{var}(X+Y) = \mathrm{var}(X) + \mathrm{var}(Y) + 2\,\mathrm{cov}(X,Y)}$.

j. A transformação de Box-Muller define o par de variáveis (Z_1, Z_2) a partir de (U_1, U_2) de acordo com

$$\begin{cases} Z_1 = \sqrt{-2\log(U_1)}\cos(2\pi U_2) \\ Z_2 = \sqrt{-2\log(U_1)}\sin(2\pi U_2) \end{cases}$$

Mostre (analiticamente, manipulando as densidades conjuntas pertinentes) que, se U_1 e U_2 são independentes e uniformemente distribuídas em $(0, 1)$, então Z_1 e Z_2 são independentes e individualmente obedecem a distribuição normal padrão (ou seja, média nula e variância unitária).

k. Determine a densidade de probabilidade da soma de duas variáveis aleatórias independentes com distribuições exponenciais (considere a possibilidade dos parâmetros que caracterizam as exponenciais serem iguais ou distintos).

l. (Kardar - particles) A corrente $I(V)$ que passa por um diodo está relacionada à voltagem aplicada V pela eq. $I(V) = I(0)\,[\exp(eV/kT) - 1]$, onde todas as constantes têm seus significados usuais. Se o diodo está sujeito a uma voltagem aleatória descrita por uma Gaussiana de média nula e variância s^2, determine a densidade de probabilidade da corrente, seu valor mais provável, seu valor médio e indique estas duas grandezas no esboço de um gráfico da densidade.

m. (estatística, peso 2) Dada uma amostra aleatória $\{X_i\}$ (v.a.'s i.i.d.), $i = 1, \cdots, n$, a *média amostral*

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

é uma *variável aleatória* sem viés, com o mesmo valor médio de cada parcela, $\langle \bar{X} \rangle = \langle X_i \rangle \equiv \mu$. Trata-se de um estimador não enviesado da média *populacional* μ.

Como podemos obter um estimador para a *variância populacional* $\mathrm{var}(X_i) \equiv \sigma^2$? Sendo a variância o segundo momento central, a primeira ideia seria calcular, na amostra, a grandeza

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2.$$

Porém, é importante perceber que isso é impossível, pois a média populacional μ é uma constante desconhecida, não completamente determinada pela amostra. Ingenuamente, uma saída parece ser o cálculo da grandeza puramente amostral

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

Mostre que este estimador de σ^2 é enviesado, com valor médio $\frac{n-1}{n}\sigma^2$, e que, portanto, a variância populacional σ^2 deve ser estimada pela *variância amostral* não enviesada

$$\boxed{S^2 \equiv \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2.}$$

n. (computacional, **peso 3**) Desenvolva uma simulação estocástica do problema apresentado no item (f), de modo que cada dígito de uma n-sequência binária da seja sorteado da forma estipulada. Gere várias dessas sequências, calcule a fração delas onde não há 1's consecutivos e compare essa fração com a resposta de (f).

o. (computacional) Como no item acima, desenvolva uma simulação correspondente ao item (g).

p. (peso 2, computacional) Use a plataforma computacional que lhe for conveniente para gerar curvas de densidade de probabilidade de distribuições exponenciais ($\lambda = 1$) e Gaussianas ($\mu = 0$ e $\sigma = 1$). Especificamente, em cada caso, você deverá (i) escolher a quantidade N de números uniformemente distribuídos u_i em $[0, 1]$ de onde você parte, (ii) converter a sequência (u_i), $i = 1, \cdots, N$ em uma sequência transformada (x_i), $i = 1, \cdots, N$ com $x_i = -\log(u_i)$ para a exponencial e Box-Muller para a Gaussiana, (iii) escolher δ (largura de cada célula a ser populada) e M (ponto de corte a partir do qual - $|x| \geq M$ - pontos são "agregados até o infinito"), (iv) estimar a distribuição empírica e, finalmente, (v) estimar a densidade de probabilidade para δ "pequeno". Comente o papel da discretização do espaço de estados na geração do ruído nas curvas.