2^a Prova - FECD - Gabarito

Renato Assunção - DCC-UFMG

Julho de 2021

1. **5 PONTOS** Vesículas são pequenas estruturas celulares de tamanhos variados e com formato aproximadamente esférico de células. Suponha que essas esferas possuem um raio aleatório R com densidade $f_R(r) = 6r(1-r) = 6(r-r^2)$ para $r \in (0,1)$. Temos interesse em obter a distribuição de probabilidade do volume aleatório $V = 4\pi/3R^3$ induzido pelo raio R.

   - Estabeleça o intervalo de valores possíveis para o volume aleatório V.
   - Para um valor v no intervalo obtido acima, obtenha a distribuição acumulada $\mathbb{F}_V(v) = \mathbb{P}(V \leq v)$.
   - Derive a função $\mathbb{F}_V(v)$ para obter a função densidade de probabilidade $f_V(v)$.
   - A densidade $f_R(r)$ do raio é mais concentrada em torno do ponto $r = 1/2$, o centro do intervalo $(0,1)$ onde os raios podem variar. A densidade $f_V(v)$ do volume também é mais concentrada em torno do ponto médio do intervalo de valores possíveis do volume? Ou ela é mais concentrada em alguma outra região desse intervalo?

   Solução:

   - O intervalo de valores possíveis para o volume aleatório V é $(0, 4\pi/3)$.
   - Para $v \in (0, 4\pi/3)$ temos

$$
\begin{aligned}
\mathbb{F}_V(v) &= \mathbb{P}(V \leq v) \\
&= \mathbb{P}(4\pi R^3/3 \leq v) \\
&= \mathbb{P}(R \leq \sqrt[3]{3v/(4\pi)}) \\
&= \int_0^{\sqrt[3]{3v/(4\pi)}} 6(r-r^2)dr \\
&= 6\left(\frac{r^2}{2} - \frac{r^3}{3}\right)\Bigg|_0^{\sqrt[3]{3v/(4\pi)}} \\
&= \frac{6}{2}\left(\frac{3v}{4\pi}\right)^{2/3} - \frac{3v}{2\pi}
\end{aligned}
$$

   - A função densidade de probabilidade $f_V(v)$ é a derivada da função acima. Para $v \in (0, 4\pi/3)$ temos

$$
f_V(v) = \left(\frac{3}{2\pi}\right)\left[\sqrt[3]{\frac{4\pi}{3v}} - 1\right]
$$

   - A densidade $f_V(v)$ é bem mais concentrada perto da origem $r = 0$. Veja que a densidade decai com $\sqrt[3]{v}$.

---

2. **5 PONTOS** Temos interesse em gerar uma amostra pelo método Monte Carlo de uma v.a. X que possui densidade de probabilidade

$$
f(x) = \begin{cases} 3x^2/8, & \text{se } x \in (0,2) \\ 0, & \text{se } x < 0 \text{ ou } x > 2 \end{cases}
$$

   Escreva uma pseudo-código (ou script python ou R) para gerar uma amostra de tamanho B usando um gerador de uma $U(0,1)$ e o

- método da acumulada inversa de Stan Ulam.
- método de aceitação e rejeição de von Neumann. Veja que você deve usar uma distribuição com densidade $g(x)$ que tenha um suporte $\mathcal{S}_g$ que **contenha** o suporte $\mathcal{S}_f$ de $f(x)$ (que é o intervalo $(0,2)$). Os suportes $\mathcal{S}_g$ e $\mathcal{S}_f$ não precisam ser idênticos mas apenas $\mathcal{S}_g \subset \mathcal{S}_f$.
- No método de aceitação e rejeição, para gerar a amostra de tamanho B de f, quantos elementos em média de g devem ser gerados?
- Use o método de amostragem por importância para estimar $\theta_1 = \mathbb{E}_f(X)$ e $\theta_2 = \mathbb{P}_f(X < 1)$. Observe que é fácil obter θ_1 e θ_2 diretamente: $\theta_1 = 3/2$ e $\theta_2 = 1/8 = 0.125$.

Solução: Para $x \in (0,2)$, a distribuição acumulada é dada por

$$\mathbb{F}(x) = \int_0^x \frac{3}{8} t^2 dt = \frac{x^3}{8}$$

e portanto a sua inversa $\mathbb{F}^{-1}$ é encontrada fazendo $u = \mathbb{F}(x)$ e invertendo $x = \sqrt[3]{8u} = \mathbb{F}^{-1}(u)$. Assim, para o método da acumulada inversa de Stan Ulam, o comando

```
x = ( 8* runif(B) )^(1/3)
```

gera o vetor x com a densidade desejada.

Para o método de aceitação-rejeição de von Neumann, vamos usar a densidade de uma uniforme em $(0,2)$. Para obter uma amostra de $U(0,2)$, basta gerar da $U(0,1)$ e multiplicar os valores por 2. A densidade $g(x)$ da $U(0,2)$ é constante e igual a $1/2$ para $x \in (0,2)$. Como a densidade alvo possui o máximo em $x = 2$ e igual a $f(2) = 3/8(2)^2 = 3/2$. Assim, podemos tomar $M = 3$ para garantir que $f(x)/(3g(x)) \leq 1$ para todo $x \in (0,2)$.

```
xg = 2 * runif(3*B)
prob = (3/8 * xg^2)/(3/2)
aceita = rbinom(3*B, 1, prob)
xf = xg[aceita == 1]
hist(xf)
```

Como $M = 3$, para gerar B valores devemos simular aproximadamente $M * B$ valores da densidade $g(x)$.

Para a amostragem por importância, temos

```
w = (3/8 * xg^2)/(1/2)
sum(w*xg)/(3*B)
sum(w*(xg<1))/(3*B)
```

---

3. **5 PONTOS** A Tabela abaixo mostra a distribuição conjunta do vetor aleatório discreto (X,Y). Obtenha: (a) a distribuição marginal da variável Y, (b) a distribuição condicional $(X|Y=2)$.

	$x=0$	$x=1$	$x=2$	$x=3$
$y=0$	0.1	0.2	0.05	0.15
$y=1$	0.1	0.05	0.1	0.15
$y=2$	0.05	0.0	0.0	0.05

Solução: A distribuição marginal da variável Y é obtida somando-se ao longo de cada linha da tabela:

$$\mathbb{P}(Y=0) = 0.1 + 0.2 + 0.05 + 0.15 = 0.50$$

$$\mathbb{P}(Y=1) = 0.1 + 0.05 + 0.1 + 0.15 = 0.40$$

$$\mathbb{P}(Y=2) = 0.05 + 0.0 + 0.0 + 0.05 = 0.10$$

A distribuição condicional $(X|Y=2)$ é obtida usnado-se o valor de $\mathbb{P}(Y=2)$ obtido acima:

$$\mathbb{P}(X=0|Y=2) = \frac{\mathbb{P}(X=0,Y=2)}{\mathbb{P}(Y=2)} = \frac{0.05}{0.10} = \frac{1}{2}$$

$$\mathbb{P}(X=1|Y=2) = \frac{\mathbb{P}(X=1,Y=2)}{\mathbb{P}(Y=2)} = \frac{0.00}{0.10} = 0$$

$$\mathbb{P}(X=2|Y=2) = \frac{\mathbb{P}(X=2,Y=2)}{\mathbb{P}(Y=2)} = \frac{0.00}{0.10} = 0$$

$$\mathbb{P}(X=3|Y=2) = \frac{\mathbb{P}(X=3,Y=2)}{\mathbb{P}(Y=2)} = \frac{0.05}{0.10} = \frac{1}{2}$$

---

4. **5 PONTOS** Um ponto X é escolhido com distribuição uniforme no intervalo $(0, L)$. Este ponto X particiona o intervalo $(0, L)$ em dois segmentos. Calcule a probabilidade de que a razão entre o segmento menor e o segmento maior seja menor que $1/4$. (Dica: faça o cálculo condicionando em cada uma das duas possibilidades, $X < L/2$ e $X \geq L/2$.)

   Solução: A especificação da razão R entre o segmento menor e o segmento maior depende da posição X em relação ao ponto central $L/2$. Se $X < L/2$, a razão é $R = X/(L-X)$, e se $X \geq L/2$, a razão é $R = (L-X)/X$. Assim, a probabilidade desejada é

$$\begin{aligned}
\mathbb{P}(R < 1/4) &= \mathbb{P}(R < 1/4 \cap X < L/2) + \mathbb{P}(R < 1/4 \cap X \geq L/2) && (1)\\
&= \mathbb{P}(R < 1/4|X < L/2)\mathbb{P}(X < L/2) + \mathbb{P}(R < 1/4|X \geq L/2)\mathbb{P}(X \geq L/2) && (2)\\
&= \mathbb{P}(X/(L-X) < 1/4|X < L/2)0.5 + \mathbb{P}((L-X)/X < 1/4|X \geq L/2)0.5 && (3)\\
&= 0.5\,(\mathbb{P}(X/(L-X) < 1/4|X < L/2) + \mathbb{P}((L-X)/X < 1/4|X \geq L/2)) && (4)
\end{aligned}$$

   Temos

$$\begin{aligned}
\mathbb{P}(X/(L-X) < 1/4|X < L/2) &= \mathbb{P}(X < L/4 - X/4|X < L/2) && (5)\\
&= \mathbb{P}(5X/4 < L/4|X < L/2) && (6)\\
&= \mathbb{P}(X < L/5|X < L/2) && (7)\\
&= \frac{\mathbb{P}([X < L/5] \cap [X < L/2])}{\mathbb{P}(X < L/2)} && (8)\\
&= \frac{\mathbb{P}(X < L/5)}{1/2} && (9)\\
&= (1/5)/(1/2) = 2/5 && (10)
\end{aligned}$$

   De modo semelhante obtemos $\mathbb{P}((L-X)/X < 1/4|X \geq L/2) = 2/5$ e portanto $\mathbb{P}(R < 1/4) = 2/5$.