Prova 02 - FECD-A

Thiago Martin Poppe

Questão 01)

Item 01)

- O intervalo possível para o valores do volume aleatório V será $(0, \frac{4\pi}{3})$.

Item 02)

- Pela definição do problema, teremos que $V = h(R)$, onde $h(R) = \frac{4\pi}{3} R^3$.
- Aplicando o método de inversão de h, teremos que $\mathbb{F}_V(v) = \mathbb{F}_R(h^{-1}(v))$.
- Precisamos então calcular o $\mathbb{F}_R(r)$ e a inversa de h para aplicarmos o método.

$$
\begin{aligned}
\mathbb{F}_R(r) &= \int_{-\infty}^{r} f_R(t)\, dt \\
&= \int_{0}^{r} 6(t - t^2)\, dt \\
&= 6 \int_{0}^{r} t - t^2 \, dt \\
&= 6 \left(\int_{0}^{r} t \, dt - \int_{0}^{r} t^2 \, dt \right) \\
&= 6 \left(\frac{t^2}{2} - \frac{t^3}{3} \right) \Bigg|_{0}^{r} \\
&= 6 \left(\frac{r^2}{2} - \frac{r^3}{3} \right) \\
&= 3r^2 - 2r^3
\end{aligned}
$$

$$
\begin{aligned}
h^{-1}(v) &= \sqrt[3]{\frac{3v}{4\pi}} \\
h^{-1}(v) &= \left[\frac{3v}{4\pi} \right]^{1/3}
\end{aligned}
$$

- Com isso, teremos que:

$$
\begin{aligned}
\mathbb{F}_V(v) = \mathbb{F}_R(h^{-1}(v)) \\
&= 3\left(\frac{3v}{4\pi}\right)^{2/3} - 2\left(\frac{3v}{4\pi}\right)^{3/3} \\
&= 3\left(\frac{3v}{4\pi}\right)^{2/3} - \frac{3v}{2\pi}
\end{aligned}
$$

Item 03)

$$
\begin{aligned}
f_V(v) = \frac{d}{dv}\mathbb{F}_V(v) &= \frac{d}{dv}\left(3\left(\frac{3v}{4\pi}\right)^{2/3} - \frac{3v}{2\pi}\right) \\
&= 3\frac{d}{dv}\left(\frac{3v}{4\pi}\right)^{2/3} - \frac{3}{2\pi}\frac{d}{dv}v \\
&= \frac{3}{2\pi}\left(\frac{3v}{4\pi}\right)^{-1/3} - \frac{3}{2\pi} \\
&= \frac{3}{2\pi}\left[\left(\frac{3v}{4\pi}\right)^{-1/3} - 1\right]
\end{aligned}
$$

Item 04)

- Pra respondermos essa pergunta, devemos computar o valor esperado de V através da seguinte forma:

$$
\begin{aligned}
\mathbb{E}(V) = \mathbb{E}(h(R)) &= \int_0^1 h(r) f_R(r)\, dr \\
&= \int_0^1 \frac{4\pi}{3} r^3 * 6(r - r^2)\, dr \\
&= 8\pi \int_0^1 r^4 - r^5\, dr \\
&= 8\pi \left(\frac{r^5}{5} - \frac{r^6}{6}\right)\Bigg|_0^1 \\
&= 8\pi * \frac{1}{30} \\
&= \frac{4\pi}{15}
\end{aligned}
$$

- Podemos notar que a densidade $f_V(v)$ não é concentrada em torno do centro do intervalo especificado anteriormente, mas sim mais próximo da região da esquerda do intervalo, próximo do valor 0.83.

Questão 02)

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

- Os testes de todos os códigos desenvolvidos foram feitos através da comparação da média da amostra com relação ao valor esperado teórico $\mathbb{E}(X) = \int_{-\infty}^{\infty} f(x)\,dx = 1.5$.

Item 01)

- Para utilizar o método da transformada inversa, teremos que computar $\mathbb{F}_X^{-1}(x)$ e gerar valores aleatórios que seguem uma uniforme $\mathcal{U}(0, 1)$.

$$\begin{aligned} \mathbb{F}_X(x) = \int_{-\infty}^{x} f_X(t)\,dt &= \int_0^x \frac{3t^2}{8}\,dt \\ &= \frac{t^3}{8}\Big|_0^x \\ &= \frac{x^3}{8}, \text{ para } x \in (0, 2) \end{aligned}$$

- Sendo assim, a função $\mathbb{F}_X^{-1}(x) = 2\sqrt[3]{x}$ e com isso, através do seguinte código em Python, nós somos capazes de gerar uma amostra de tamanho B para X.

```
import numpy as np

def Fx_inverse(x):
    return 2 * np.cbrt(x)

B = 10000
X_sample = Fx_inverse(np.random.uniform(0, 1, size=B))
```

Item 02)

- Para esse método usarmos a distribuição auxiliar uniforme $\mathcal{U}(0, 2)$, cuja densidade $g(x)$ possui o mesmo suporte que a densidade de X. Note que nesse caso nós não teremos um truncamento e "viciamento" da geração da amostra uma vez que para valores de $x < 0$ e $x > 2$ a função de densidade $f(x)$ assume valor 0, sendo assim exatamente o mesmo suporte da distribuição auxiliar escolhida.
- O próximo passo consiste em escolhermos uma constante M tal que $f(x) \leq Mg(x)$ para todo x, escolheremos $M = 4$.
- Definimos agora a razão entre as alturas das funções $r(x) = \frac{f(x)}{Mg(x)} < 1$ e utilizamos o seguinte código em Python para gerar a nossa amostra realizando n iterações. Note que o código abaixo **não** gera uma amostra de tamanho B... Para isso, devemos gerar em média $M * B$ valores da distribuição auxiliar, ou no caso do código realizar $M * n$ iterações!

```
import numpy as np
import scipy.stats as ss

def f(x):
    return (3 * x**2) / 8

M = 4
n_iter = 1000000

# Gerando n_iter amostras de uma U(0,2) e computando as razões
U_sample = ss.uniform(0,2).rvs(n_iter)
ratios = f(U_sample) / (M * ss.uniform(0,2).pdf(U_sample))

# Jogando uma moeda para cima e a nossa amostra será
# apenas as jogadas que resultaram em "cara" (valor 1)
coin = ss.binom(1, ratios).rvs(n_iter)
X_sample = U_sample[coin == 1]
```

Item 03)

- Como discutido no item anterior, caso quisermos gerar uma amostra de tamanho B, devemos esperar em média a geração de $M * B$ valores da distribuição auxiliar. Sendo assim, precisamos escolher um valor de M consciente para não necessitar de muitas iterações!

Item 04)

- Este método é relativamente simples, onde (i) precisamos definir uma distribuição auxiliar de forma que a sua densidade $g(x)$ contenha o suporte de $f(x)$, (ii) computar pesos $w_i = \frac{f(x_i)}{g(x_i)}$ e (iii) aplicamos tais pesos na amostra gerada a partir da distribuição auxiliar.
- O seguinte código em Python demonstra a aplicação desse método:

```
import numpy as np
import scipy.stats as ss

def f(x):
    return (3 * x**2) / 8

B = 10000

# Gerando B amostras de uma U(0,2) e computando os pesos
U_sample = ss.uniform(0,2).rvs(B)
weigths = f(U_sample) / ss.uniform(0,2).pdf(U_sample)

X_sample = weigths * U_sample
```

Questão 03)

Letra a)

- Para determinarmos a distribuição marginal de Y, precisamos computar $\mathbb{P}(Y = y)$ para $y \in \{0, 1, 2\}$. Sendo assim:
    1. $\mathbb{P}(Y = 0) = \sum_{x=0}^{3} \mathbb{P}(Y = 0 \wedge X = x) = 0.1 + 0.2 + 0.05 + 0.15 = 0.5$
    2. $\mathbb{P}(Y = 1) = \sum_{x=0}^{3} \mathbb{P}(Y = 1 \wedge X = x) = 0.1 + 0.05 + 0.1 + 0.15 = 0.4$
    3. $\mathbb{P}(Y = 2) = \sum_{x=0}^{3} \mathbb{P}(Y = 2 \wedge X = x) = 0.05 + 0.00 + 0.00 + 0.05 = 0.1$

Letra b)

- Para computar a distribuição condicional $(X \mid Y = 2)$ nós devemos eliminar da tabela todas as entradas onde $Y \neq 2$, ficando assim apenas com os valores do "mundo" onde $Y = 2$.

- Como queremos a distribuição, devemos normalizar esses valores para que a soma das novas probabilidades seja igual à 1. Sendo assim, teremos a seguinte tabela:

X	Y	$\mathbb{P}(X\|Y=2)$
0	2	0.5
1	2	0.0
2	2	0.0
3	2	0.5

Questão 04)

- Seja r a razão entre o menor segmento e o maior segmento, queremos saber o valor de $\mathbb{P}(r < \frac{1}{4})$. Para computar esse valor iremos condicionar a resposta em dois possíveis casos.

1. Caso onde $X < \frac{L}{2}$.
   - Para esse caso, teremos que $r = \frac{|menor|}{|maior|} = \frac{X}{L-X}$.
   - Note que o valor da razão é uma variável aleatória induzida pelo ponto X. Sendo assim, iremos modelar $Y = h_1(X) = \frac{X}{L-X}$.
   - Com isso, agora o problema passa a ser $\mathbb{P}(Y < \frac{1}{4}) = \mathbb{F}_Y(y = \frac{1}{4})$. Usaremos o método da transformada inversa para encontrar tal resultado.
   - Sabemos que $\mathbb{F}_X(x) = \frac{x}{L}$, já que $X \sim U(0, L)$, e conseguimos calcular que a inversa de h_1 será $\frac{yL}{1+y}$.
   - Com isso, podemos encontrar que $\mathbb{F}_Y(y) = \mathbb{F}_X(h_1^{-1}(y)) = \frac{y}{1+y}$.
   - Finalmente, teremos que $\mathbb{P}(r < \frac{1}{4}) = \mathbb{P}(Y < \frac{1}{4}) = \mathbb{F}_Y(\frac{1}{4}) = \frac{1}{5}$.
2. Caso onde $X \geq \frac{L}{2}$.
   - Para esse caso, teremos que $r = \frac{|menor|}{|maior|} = \frac{L-X}{X}$.
   - Note que o valor da razão é uma variável aleatória induzida pelo ponto X. Sendo assim, iremos modelar $Z = h_2(X) = \frac{L-X}{X}$.
   - Com isso, agora o problema passa a ser $\mathbb{P}(Z < \frac{1}{4}) = \mathbb{F}_Z(z = \frac{1}{4})$. Usaremos o método da transformada inversa para encontrar tal resultados.

- Sabemos que $\mathbb{F}_X(x) = \frac{x}{L}$, já que $X \sim U(0, L)$, e conseguimos calcular que a inversa de h_2 será $\frac{L}{1+z}$.
- Com isso, podemos encontrar que $\mathbb{F}_Z(z) = \mathbb{F}_X(h_2^{-1}(z)) = \frac{1}{1+z}$.
- Finalmente, teremos que $\mathbb{P}(r < \frac{1}{4}) = \mathbb{P}(Z < \frac{1}{4}) = \mathbb{F}_Z(\frac{1}{4}) = \frac{4}{5}$.