UFRJ - CCMN - IM - Departamento de Métodos Estatísticos

Questões de simulação - turma MAI - entrega até o dia 29 de novembro

---

1. Sejam X e Y variáveis aleatórias independentes e identicamente distribuídas segundo uma distribuição Exponencial(1) e defina as seguintes variáveis aleatórias $Z = X + Y$, $W = \max\{X, Y\}$ e $V = \min\{X, Y\}$.

   (a) Determine as densidades das variáveis aleatórias Z, W e V bem como seus respectivos valores esperados e variâncias.

   (b) Escreva um programa que simule as variáveis X e Y e, a partir desta simulação, geram valores das demais variáveis Z, W e Y.

   (c) Gere 10.000 valores de cada uma das variáveis e construa histogramas dos valores gerados.

   (d) Compare os histogramas obtidos com as densidades teóricas. Sugestão: Construa os histogramas na escala da densidade de frequência relativa.

2. Seja X uma variável aleatória com a seguinte função de distribuição.

$$F(x) = \begin{cases} 0, & x < 0 \\ x^3, & 0 \leq x \leq 1 \\ 1, & x > 1 \end{cases}$$

   (a) Calcule o valor esperado e a variância de X.

   (b) Escreva um programa que gere valores dessa variável aleatória.

   (c) Use o programa para gerar 10.000 valores dessa variável aleatória.

   (d) Construa um histograma dos valores gerados e compare este histograma com a forma da densidade da variável aleatória X. Sugestão: Construa o histograma na escala da densidade de frequência relativa.

   (e) Usando os valores gerados de X construa um intervalo de 95% de confiança para o valor esperado de X.

3. Sejam $X_1, X_2, \cdots$ variáveis aleatórias independentes e identicamente distribuídas, segundo a distribuição Uniforme em (0,1). Defina

$$N = \min\left\{n : \sum_{i=1}^{n} X_i > 1\right\}$$

   Isto é, N é igual ao menor número de variáveis aleatórias X_i's cuja soma é maior do que 1.

   (a) Estime $\mathsf{E}[N]$, gerando 100 valores de N.

   (b) Estime $\mathsf{E}[N]$, gerando 1.000 valores de N.

   (c) Estime $\mathsf{E}[N]$, gerando 10.000 valores de N.

   (d) Gere 100 amostras de tamanho 1000 da variável N. Para cada uma delas calcule a média observada. Depois tome os quantis de 2,5% e de 97,5% das médias obtidas como um intervalo de 95% de confinaça para $\mathsf{E}[N]$.

(e) Gere uma amostra de tamanho 1000 da variável N e use o teorema central do limite para construir um intervalo de 95% de confiança para média. Substitua no lugar do desvio padrão de N sua estimativa baseada na amostra de tamanho 1000.

4. Sejam F_i, $i = 1, 2, \cdots k$ funções de distribuição.

   (a) Mostre que $F(x) = \sum_{i=1}^{k} a_i \cdot F_i(x)$, com $0 < a_i < 1$, para todo $i \in \{1, 2, \cdots k\}$ e $\sum_{i=1}^{k} a_i = 1$, é uma função de distribuição. Dizemos que $F(x)$ é uma mistura das distribuições F_1, F_2, $\cdots$, F_k.

   (b) Explique como você pode gerar valores de uma variável cuja distribuição é uma mistura como a do item (a), conhecendo geradores das distribuições F_i, $i = 1, 2, \cdots n$.

   (c) Use o item anterior para gerar valores de uma variável aleatória cuja função de distribuição é dada por

$$F(x) = \begin{cases} 0, & x \leq 0 \\ \frac{x + 2x^3 + x^5}{4}, & 0 < x < 1 \\ 1, & x \geq 1 \end{cases}$$

   (d) Gere 1000 valores dessa variável, construa um histograma dos valores obtidos e compare a forma do histohrama obtido com a forma da densidade de probabilidade dessa variável aleatória.

5. Clientes chegam ao acaso em uma estação de serviço na qual há apenas um atendente e cada cliente é atendido imediatamente se não há nenhum cliente na sua frente, mas ele deve esperar em fila caso contrário. O problema a ser resolvido é calcular o tempo médio de espera de um cliente, segundo as seguintes suposições:

   - a estação de serviço permite a entrada de clientes a partir de 9h até às 17h, ou seja, durante um período de 8 horas;
   - os tempos entre chegadas sucessivas de clientes são independentes e identicamente distribuídos segundo uma distribuição exponencial de média μ;
   - os tempos de atendimento de cada cliente são variáveis aleatórias independentes e identicamente distribuídas segundo uma distribuição exponencial de média λ.
   - os tempos entre chegadas e os tempos de atendimento são independentes.

   Observe que o tempo de espera de um cliente para ser atendido dependerá de quantos clientes estiverem na frente dele e quanto tempo levará para que os clientes a sua frente sejam atendidos. Além disso, após às 17 horas, se houver clientes na fila, os mesmos serão atendidos quando chegar a sua vez, mas não será permitida a entrada de novos clientes.

   (a) Escreva um programa que simule o comportamento de um dia de serviço supondo que o tempo médio entre duas chegadas sucessivas seja $\mu = 10$ minutos e o tempo médio de atendimento seja $\lambda = 5$ minutos. Registre o número total de clientes do dia e o tempo médio de espera.

   (b) Calcule o número esperado e o desvio padrão do número de clientes em um dia de serviço.

   (c) Execute o programa elaborado no item (a) $N = 100$ vezes e construa um intervalo de confiança de 90% para o tempo médio de espera na fila de um cliente em um dia nessa estação.

(d) Comparando com a situação simulada, o que você espera que irá ocorrer com o tempo médio de espera se fizermos $\mu = 5$ minutos e $\lambda = 10$ minutos? Justifique sua resposta, construindo um intervalo de confiança para o tempo médio de espera na fila de 90% de confiança sob essa configuração.

6. Suponha que deseja-se testar se um dado é honesto. Lançamos o dado n vezes e registramos o número de faces 1, X_1, o número de faces 2, X_2, e, assim por diante, até o númerode faces 6, X_6 obtidas. Defina $X = \sum_{i=1}^{6} \frac{(X_i - n/6)^2}{n/6}$. Para um dado honesto, a distribuição da variável X é uma qui-quadrado com 5 graus de liberdade (equivalente a uma Gamma com parâmetros 5/2 e 1/2).

   (a) Fazendo $n = 60$, simule o experimento 1000 vezes, considerando o dado honesto, e para cada simulação guarde o valor de X.

   (b) Construa um histograma dos valores de X e compare com a distribuição de qui-quadrado com 5 graus de liberdade. Na construção do histograma, use a escala da densidade de frequência.

   (c) Repita o item (a), mas agora tornando o dado desonesto de tal modo que os números ímpares sejam equiprováveis, os números pares também sejam equiprováveis, mas os números ímpares ocorrem com probabilidade duas vezes maior do que os números pares. Depois compare o histograma dos valores obtidos de X com a densidade de qui-quadrado com 5 graus de liberdade.