

Lista 06 de Exercícios FECD

V.A.s, Transformações e Testes de aderência

Renato Assunção - DCC, UFMG

2018

Exercícios

Nesta lista, temos muitos exercícios para seu treinamento e prazer. Você deve entregar APENAS 5 exercícios. Você pode escolhê-los à vontade.

1. Os primeiros 608 dígitos da expansão decimal do número π tem as seguintes frequências:

k	0	1	2	3	4	5	6	7	8	9
Obs	60	62	67	68	64	56	62	44	58	67
Esp	??	??	??	??	??	??	??	??	??	??

Use o teste qui-quadrado para responder se estes dados são compatíveis com a suposição de que cada dígito é escolhido de forma completamente aleatória. Isto é, de acordo com uma distribuição uniforme discreta sobre os possíveis dígitos.

2. O R possui uma função, `ks.test()`, que implementa o teste de Kolmogorov. Suponha que x é um vetor com n valores numéricos distintos. Então `ks.test(x,"pnorm",m,dp)` testa se x pode vir da distribuição $N(m, dp)$, uma normal (ou gaussiana) com média $\mu = m$ e desvio-padrão $\sigma = dp$. Outras distribuições são possíveis substituindo o string "pnorm": as pré-definidas em R (veja com `?distributions`) ou qualquer outra para a qual você crie uma função que calcula a função distribuição acumulada teórica.

```
> ks.test(x, "pnorm")
```

One-sample Kolmogorov-Smirnov test

```
data: x
D = 0.0805, p-value = 0.876
alternative hypothesis: two-sided
```

A saída de `ks.test()` fornece o valor de $D_n = \max_x |\hat{F}_n(x) - F(x)|$ e o seu p-valor. Dissemos em sala que se $\sqrt{n}D_n > 1.36$, rejeitamos o modelo. Caso contrário, não há muita evidência nos dados para rejeitar o modelo (não quer dizer que o modelo seja correto, apenas não conseguimos rejeitá-lo).

Gere alguns dados com $n = 50$ de uma normal qualquer e use a função `ks.test()` para verificar se o teste rejeita o modelo. Faça o teste de dois modos: use o modelo correto que você usou para gerar seus dados e depois use um modelo diferente deste alterando, por exemplo, o valor de μ ou σ .

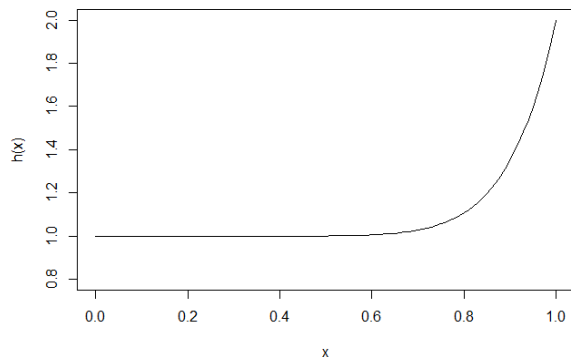


Figura 1: Gráfico da função $h(x)$ usada para criar a v.a. $Y = h(X)$ onde $X \sim \text{Unif}(0, 1)$.

3. Implemente em R uma função para calcular o resultado de um teste de Kolmogorov. A função estará restrita a testar apenas o modelo normal com média $\mu = \text{m}$ e desvio-padrão $\sigma = \text{dp}$ que devem ser fornecidas pelo usuário ou obtidas dos próprios dados (default) usando a média aritmética (comando `mean()`) e o desvio-padrão amostral (raiz da saída do comando `var()`). Não se preocupe em lidar com os casos extremos (usuário fornecer vetor nulo, fornecer vetor com valores repetidos, etc).

Observação importante: pode-se provar que para encontrar $D_n = \max_x |\hat{F}_n(x) - F(x)|$ basta varrer os pontos de salto de $\hat{F}_n(x)$, olhando o valor de $\hat{F}_n(x)$ imediatamente antes de x_i ou no próprio ponto x_i , onde x_i é um dos valores do vetor de dados observados.

4. Seja $Y = h(X)$ onde $X \sim \text{Unif}(0, 1)$. A função $h(x)$ é mostrada no gráfico da Figura 1. A partir dessa figura, é possível obter aproximadamente os valores da f.d.a. $\mathbb{F}_Y(y)$ sem fazer nenhum cálculo explícito, apenas no olhômetro. Dentre as opções abaixo, decida qual o valor que melhor aproxima $\mathbb{F}_Y(y)$.

- $\mathbb{F}_Y(0.9)$ é aproximadamente igual a: (a) zero, (b) 0.01 (c) 0.5 (d) 0.8 (e) 0.95 (f) um.
- $\mathbb{F}_Y(1.1)$ é aproximadamente igual a: (a) zero, (b) 0.01 (c) 0.5 (d) 0.8 (e) 0.95 (f) um.
- $\mathbb{F}_Y(1.8)$ é aproximadamente igual a: (a) zero, (b) 0.01 (c) 0.5 (d) 0.8 (e) 0.95 (f) um.
- $\mathbb{F}_Y(2.1)$ é aproximadamente igual a: (a) zero, (b) 0.01 (c) 0.5 (d) 0.8 (e) 0.95 (f) um.

5. Transformação de v.a.'s: Seja X o lado de um quadrado aleatório. A v.a. X é selecionada de uma distribuição $\text{Unif}(0, 1)$. A área do quadrado formado com lado X é a v.a. $Y = X^2$.

- Calcule o comprimento esperado do lado do quadrado $\mathbb{E}(X)$.
- Obtenha também a área esperada $\mathbb{E}(Y)$. É verdade que $\mathbb{E}(Y) = (\mathbb{E}(X))^2$? Ou seja, a área esperada $\mathbb{E}(Y)$ é igual à $(\mathbb{E}(X))^2$, a área de um quadrado cujo lado tem comprimento igual ao comprimento esperado?
- Qual a distribuição de Y ? Isto é, obtenha $F_Y(y)$ para $y \in \mathbb{R}$. item Derive $F_Y(y)$ para obter a densidade $f_Y(y)$ e faça seu gráfico. Qual a região onde mais massa de probabilidade é alocada? O que é mais provável, um quadrado com área menor que 0.1 ou maior que 0.9?

-
6. Refaça o exercício anterior considerando o volume aleatório $V = X^3$ do cubo aleatório formado com o lado $X \sim \text{Unif}(0, 1)$.
-

7. Refaça o exercício anterior considerando o volume aleatório $V = (4/3)\pi X^3$ da esfera aleatória formada com o raio $X \sim \text{Unif}(0, 1)$.
-

8. Para este exercício, veja texto no final da lista. Supondo X contínua com densidade $f(x)$ e S um subconjunto da reta real. Complete os passos da dedução abaixo preenchendo os locais indicados por ??.

$$\begin{aligned}\mathbb{P}(X \in S) &= \int_{??} f(x) dx \\ &= \int_{??} I_S(x) f(x) dx \\ &= \mathbb{E}(I_{??}(??))\end{aligned}$$

9. Seja $Y = h(X) = 1 + X^{10}$ onde $X \sim U(0, 1)$, uma uniforme no intervalo real $(0, 1)$. Isto é, a probabilidade de que X caia num intervalo (a, b) contido em $(0, 1)$ é o comprimento $b - a$ do intervalo. A Figura 1 mostra o gráfico desta transformação.

- Os valores possíveis de Y formam o intervalo $(?, ?)$. Complete os locais marcados com “?”.
 - Analisando a Figura 1 verifique aonde o intervalo $(0, 0.8)$ no eixo x é levado pela transformação no eixo $y = h(x)$. Faça o mesmo com o intervalo de mesmo comprimento $(0.8, 1)$. Conclua: $\mathbb{P}(Y \in (1.2, 2.0))$ é maior ou menor que $\mathbb{P}(Y \in (0, 1.1))$?
 - Sejam os eventos $B = [Y < ?]$ e $A = [X < 1/\sqrt[10]{2}]$, onde $1/\sqrt[10]{2} \approx 0.933$. Os eventos A e B devem ser são iguais. Qual o valor de “?”?
 - Considerando a distribuição de Y , calcule $F_Y(y) = \mathbb{P}(Y \leq y)$ para qualquer $y \in \mathbb{R}$ mapeando o evento $[Y \leq 1/2]$ e um evento equivalente $[X \in S]$ e calculando $\mathbb{P}(X \in S)$.
 - Derive $F_Y(y)$ para obter a densidade $f(y)$ de Y . Esboce a densidade e com base no gráfico, sem fazer contas, responda: o que é maior, $\mathbb{P}(Y < 1/2)$ ou $\mathbb{P}(Y > 1/2)$?
-

10. Considere a f.d.a. $\mathbb{F}(x) = \mathbb{P}(X \leq x)$. Quais afirmações abaixo são corretas?

- $\mathbb{F}(x)$ é uma função aleatória.
 - Se X é uma v.a. discreta então $\mathbb{F}(x)$ possui saltos em todos os pontos onde X tem massa de probabilidade maior que zero.
 - $\mathbb{F}(x)$ mede a probabilidade de X ser menor que média.
 - $\mathbb{F}(x)$ é uma função determinística.
 - $\mathbb{F}(x)$ só pode ser calculada depois que uma amostra é obtida.
 - $\mathbb{F}(x)$ é a mesma função, qualquer que seja a amostra aleatória de X .
-

11. Exercício para verificar aprendizagem de notação: Seja X_1, X_2, \dots, X_n uma amostra de uma v.a. Considere a f.d.a. empírica

$$\hat{\mathbb{F}}(x) = \frac{1}{n} \quad (\text{no. elementos} \leq x)$$

Explique por que isto é equivalente a escrever

$$\hat{\mathbb{F}}(x) = \frac{\sum_{i=1}^n I_{[X_i \leq x]}}{n} = \frac{\sum_{i=1}^n I_{(-\infty, x]}(X_i)}{n}$$

12. Considere a f.d.a. empírica $\hat{\mathbb{F}}(x) = \sum_i I[X_i \leq x]/n$ baseada numa amostra aleatória de X . Quais afirmações abaixo são corretas?

- $\hat{\mathbb{F}}(x)$ é uma função aleatória.
- Se X é uma v.a. discreta então $\hat{\mathbb{F}}(x)$ possui saltos em todos os pontos onde X tem massa de probabilidade maior que zero.
- $\hat{\mathbb{F}}(x)$ mede a probabilidade de X ser menor que média.
- $\hat{\mathbb{F}}(x)$ é uma função determinística.
- $\hat{\mathbb{F}}(x)$ só pode ser calculada depois que uma amostra é obtida.
- $\hat{\mathbb{F}}(x)$ é a mesma, qualquer que seja a amostra aleatória de X .
- $\mathbb{F}(x)$ é a mesma função, qualquer que seja a amostra aleatória de X .

13. Em finanças, o valor presente (hoje) de um capital c a ser pago daqui a T anos é dado por $V = c \exp(-\delta T)$ onde δ é a taxa de juros anual. Um valor típico é $\delta = 0.04$, o que corresponde a 4% anuais de juros. Imagine que c é o capital a ser pago por uma apólice de seguros a um beneficiário quando um indivíduo falecer. Se T é o tempo de vida futuro (e aleatório) deste indivíduo, $V = c \exp(-\delta T)$ representa o valor atual (presente, no instante da assinatura do contrato da apólice) deste capital futuro e incerto. Para precificar o seguro e estabelecer o prêmio a ser cobrado do segurado, a seguradora precisa calcular o valor esperado $\mathbb{E}(V)$. Supondo que T possui uma distribuição exponencial com parâmetro $\lambda = 1/40$ (ou média igual a 40), obtenha $\mathbb{E}(V)$. OBS: a densidade de uma exponencial com parâmetro λ é dada por

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

14. Suponha que X é uma variável aleatória assumindo valores no intervalo $(0, 1)$ e que $Y = X^2$. Isto é, Y é a área de um quadrado de lado aleatório X . A v.a. X NÃO segue uma uniforme $U(0, 1)$. Na verdade, deseja-se que a área Y é que tenha uma distribuição uniforme $U(0, 1)$. Qual deve ser a densidade $f_X(x)$ do lado aleatório X de tal forma que $Y \sim U(0, 1)$?

15. Uma seguradora possui uma carteira com 50 mil apólices de seguro de vida. Não é possível prever quanto cada pessoa vai viver mas é possível prever o comportamento estatístico dessa massa de segurados. Atuários estudam este fenômeno e já identificaram uma distribuição excelente para o tempo de vida X de indivíduos a partir de seu nascimento: a distribuição de Gompertz que possui densidade de probabilidade $f_X(x)$ dada por:

$$f(x) = Bc^x \exp\left(-\frac{B}{\log(c)}(c^x - 1)\right) = Bc^x S(x)$$

para $x \geq 0$, onde $B > 0$ e $c \geq 1$ são constantes positivas que alteram o formato da função densidade. Evidentemente, $f(x) = 0$ para $x < 0$ pois não existe tempo de vida negativo. A função de distribuição acumulada é igual a

$$\mathbb{F}_X(x) = 1 - e^{-\frac{B(c^x - 1)}{\log(c)}}$$

para $x \geq 0$ e $\mathbb{F}_X(x) = 0$ para $x < 0$. Usando dados recentes de uma seguradora brasileira, podemos tomar $B = 1.02 \times 10^{-4}$ e $c = 1.0855$.

- Com os parâmetros B e c acima, desenhe a curva densidade de probabilidade (use valores x entre 0 e 100 anos).
- SEM FAZER NENHUMA conta, apenas olhando a curva que você gerou, responda:
 - $\mathbb{P}(X < 40)$ é aproximadamente igual a 0.03, 0.10, ou 0.20?
 - Deslize mentalmente um pequeno intervalo de um ano e considere todas as probabilidades do tipo $\mathbb{P}(X \in [k, k + 1))$ onde k é um natural. Qual a idade em que esta probabilidade é aproximadamente máxima: aos $k = 60, 70$ ou 80 anos de idade?
 - O que é maior, a probabilidade de morrer com mais de 100 anos ou de morrer antes de completar 10 anos de idade?
- Inverta a função de distribuição acumulada, mostrando que

$$F^{-1}(u) = \log(1 - \log(c) \log(1 - u)/B) / \log(c)$$

onde $u \in (0, 1)$.

- Use o método da transformada inversa para gerar 50 mil valores independentes de tempos de vida X . Este método é muito simples. Basta gerar $U \sim U(0, 1)$ e a seguir transformar gerando $Y = F^{-1}(U)$. Então Y segue uma Gompertz com parâmetros B e c .
- Com estes números simulados, calcule aproximadamente $\mathbb{P}(X > 80 | X > 30)$. Isto é, calcule aproximadamente a chance de sobreviver pelo menos mais 50 anos dado que chegou a completar 30 anos de idade.
- A seguradora cobra um prêmio de 2 mil reais por uma apólice de seguro de vida que promete pagar 100 mil reais a um beneficiário no momento exato de morte do segurado. A apólice é vendida no momento em que os 50 mil indivíduos nasceram (alterar esta hipótese para que apenas adultos comprem a apólice dá trabalho e não muda o essencial do exercício). Ela coloca o dinheiro rendendo juros de 5% ao ano de forma que dentro de t anos os 2 mil reais terão se transformado em $2 \times \exp(0.05t)$. Se o indivíduo falecer muito cedo, ela terá uma perda financeira. Se ele sobreviver muito tempo, seu prêmio vai acumular juros suficiente para cobrir o pagamento do benefício.

Para a carteira de 50 mil vidas que você gerou, calcule aproximadamente a probabilidade de perda financeira da seguradora. A perda é dada por $L = 2 \sum_i \exp(0.05t_i) - (100 \times 50000)$ e o índice i varia sobre os indivíduos. Você deseja calcular aproximadamente $\mathbb{P}(L < 0)$. Para isto, simule toda a carteira de 50 mil vidas umas 100 vezes e calcule a proporção das vezes em que L foi menor que zero.

Probabilidades e variáveis aleatórias indicadoras

Toda probabilidade pode ser vista como a esperança de uma v.a. indicadora. Seja A um evento qualquer e $I_A(\omega)$ uma v.a. binária dada por

$$I_A(\omega) = \begin{cases} 1, & \text{se } \omega \in A \\ 0, & \text{caso contrário} \end{cases}$$

A função é apenas a função indicadora de que o evento A ocorreu. Para cada resultado do experimento ela retorna 1 se A ocorreu e 0 se A não ocorreu.

O evento $[I_A = 1]$ é o conjunto de elementos $\omega \in \Omega$ tais que $I_A(\omega) = 1$. Mas isto é exatamente o conjunto de $\omega \in A$. Ou seja, temos a igualdade $[I_A = 1] = A$ de dois eventos em Ω . Portanto, a probabilidade dos dois eventos é a mesma e $\mathbb{P}(I_A = 1) = \mathbb{P}(A)$.

Como a v.a. I_A é binária temos

$$\mathbb{E}(I_A) = 1 \times \mathbb{P}(I_A = 1) + 0 \times \mathbb{P}(I_A = 0) = \mathbb{P}(I_A = 1) = \mathbb{P}(A)$$

Assim, para todo evento A , podemos escrever sua probabilidade $\mathbb{P}(A)$ como a esperança $\mathbb{E}(I_A)$ da v.a. aleatória indicadora da ocorrência de A . Parece um jeito muito complicado de obter uma probabilidade mas este é um truque muito útil em situações um pouco mais complicadas. A utilidade vem da possibilidade de usar algumas propriedades conhecidas do operador \mathbb{E} , e isto torna às vezes mais fácil de manipular que o operador \mathbb{P} . Uma dessas propriedades é a linearidade da esperança: $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.

Eventos podem ser definidos em termos de variáveis aleatórias. Por exemplo,

- $[X = 3] = \{\omega \in \Omega \text{ tais que } X(\omega) = 3\}$
- $[X \leq 3] = \{\omega \in \Omega \text{ tais que } X(\omega) \leq 3\}$
- $[1 \leq X \leq 3] = \{\omega \in \Omega \text{ tais que } X(\omega) \in [1, 3]\}$
- se $S \subset \mathbb{R}$ temos $[X \in S] = \{\omega \in \Omega \text{ tais que } X(\omega) \in S\}$

Assim, podemos estender a ideia de escrever como esperança a probabilidade de um evento associado a uma v.a. Por exemplo, seja A o evento $A = [X = 3]$ associado com a v.a. X . Então

$$\mathbb{P}(X = 3) = \mathbb{P}(A) = \mathbb{E}(I_A) = \mathbb{E}(I_{[X=3]})$$

De maneira geral, se $S \subset \mathbb{R}$ temos $\mathbb{P}(X \in S) = \mathbb{E}(I_{[X \in S]})$.

Vamos adotar outra notação alternativa:

$$I_{[X \in S]} = I_S(X)$$

Veja que $I_S(X)$ é uma v.a. Ela recebe como input um resultado $\omega \in \Omega$ e retorna um valor binário $I_{[X \in S]}(\omega) = I_S(X(\omega))$. Esta v.a. é uma transformação $h(X)$ de X . De fato, temos

$$I_S(X(\omega)) = h(X(\omega)) = \begin{cases} 1, & \text{se } X(\omega) \in S \\ 0, & \text{caso contrário} \end{cases}$$

Portanto, se tivermos a densidade $f(x)$ de uma v.a. contínua X , usando o que aprendemos no início deste texto, podemos escrever

$$\begin{aligned} \mathbb{P}(X \in S) &= \mathbb{E}(I_S(X)) \\ &= \mathbb{E}(h(X)) \\ &= \int_{\mathbb{R}} h(x)f(x)dx \\ &= \int_{\mathbb{R}} 1_S(x)f(x)dx \\ &= \int_S 1 \times f(x)dx + \int_{\mathbb{R}-S} 0 \times f(x)dx \\ &= \int_S f(x)dx \end{aligned}$$

Eu sei que parece estarmos dando voltas em torno do mesmo ponto mas, acredite, isto é útil. Por exemplo, suponha que $X \sim N(0, 1)$, uma gaussiana padrão que possui densidade

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right)$$

Suponha que, por alguma razão queiramos calcular a probabilidade de que $\exp(-|X|) > |X|$. Manipulação algébrica mostra que $\exp(-|x|) > x$ se, e somente se, $x \in (-a, a)$ onde $a \approx 0.5671$. Assim,

$$\mathbb{P}(\exp(-|X|) > |X|) = \int_{(-a, a)} f(x) dx = \int_{\mathbb{R}} h(x) f(x) dx = \mathbb{E}(h(X))$$

onde $h(X) = I_S(X)$ e S é o evento $[\exp(-|X|) > |X|]$.