Fundação CECIERJ - **Vice Presidência de Educação Superior a Distância**

Curso de Tecnologia em Sistemas de Computação
Disciplina Probabilidade e Estatística
AD2 1° semestre de 2014

1 - Primeira questão (2,5 pontos)

Verifique quais das funções abaixo podem ser usadas como densidades de probabilidade dentro dos intervalos especificados. Se o que inviabiliza o uso for que a função não seja normalizada, normalize e apresente a distribuição obtida. Determine também a variância das distribuições de probabilidade encontradas que não necessitaram de normalização.

Resolução:

Para serem densidades de probabilidade é necessário que a função seja não negativa dentro do intervalo em que é definida e que a soma de todas as probabilidades seja 1, ou seja, formalmente se escreve

$$\int_{-\infty}^{\infty} f(x)\,dx = 1 \quad .$$

$a.\ f(x)=3x^2\ ,\ se\ -1 \le x \le 1.$

Resolução:

Observemos que a função é não negativa em todo o intervalo especificado. Calculemos a integral da mesma

$$\int_{-1}^{1} f(x)\,dx = \int_{-1}^{1} 3x^2\,dx = 3\int_{-1}^{1} x^2\,dx = \frac{3}{3}x^3\Big|_{-1}^{1} = (1+1) = 2 \quad .$$

Para que esta função seja densidade de probabilidade será necessário normalizá-la, onde 2 é o fator de normalização. Assim teremos que a função abaixo é uma densidade de probabilidade

$$f(x) = \frac{3}{2}x^2 \quad .$$

$b.\ f(x)=x-e^x\ ,\ se\ -1 \le x \le 1.$

Resolução:

É fácil de observar que a função acima toma valores negativos dentro do intervalo especificado. Por exemplo, em x = 0 função toma o valor -1. Portanto, esta função não é distribuição de probabilidade.

$c.\ f(x)=x^3-x\ ,\ se\ 0 \le x \le 2.$

Resolução:

Por simples inspeção verificamos que esta função é negativa no intervalo [0, 1). Portanto, ela não é uma distribuição de probabilidade.

$d. \quad f(x)=x \ , \ se \ \ 0\leq \ x \ \leq\sqrt{2}.$

Resolução:

Claramente a função é não negativa no intervalo dado. Passemos a integração.

$$\int_0^{\sqrt{2}} f(x)\,dx=\int_0^{\sqrt{2}} x\,dx=\frac{x^2}{2}\Big|_0^{\sqrt{2}}=\frac{2}{2}=1 \quad .$$

ou seja, esta é uma distribuição de probabilidades. Sendo assim, calculemos a variância. Esta é dada pela expressão

$$\sigma^2=\int_a^b x^2 f(x)\,dx-\mu^2$$

e, por sua vez, $\mu=\int_a^b x f(x)\,dx$ **.**

Calculemos a média.

$$\mu=\int_0^{\sqrt{2}} x f(x)\,dx=\int_0^{\sqrt{2}} x\,x\,dx=\int_0^{\sqrt{2}} x^2\,dx=\frac{x^3}{3}\Big|_0^{\sqrt{2}}=\frac{(\sqrt{2})^3}{3}=\frac{2}{3}\sqrt{2} \quad .$$

Calculemos agora a integral abaixo

$$\int_0^{\sqrt{2}} x^2 f(x)\,dx=\int_0^{\sqrt{2}} x^2\,x\,dx=\int_0^{\sqrt{2}} x^3\,dx=\frac{x^4}{4}\Big|_0^{\sqrt{2}}=\frac{(\sqrt{2})^4}{4}=\frac{4}{4}=1 \quad .$$

Compondo os dois resultados teremos para a variância

$$\sigma^2=\int_a^b x^2 f(x)\,dx-\mu^2=1-\left(\frac{2}{3}\sqrt{2}\right)^2=1-\frac{4}{9}\times 2=1-\frac{8}{9}=\frac{1}{9} \quad .$$

$e. \quad f(x)= \begin{matrix} (2+x)/4, \ se \ -2\leq \ x \ \leq 0; \\ (2-x)/4, \ se \ 0\leq \ x \ \leq 2. \end{matrix}$

Resolução:

Note que no primeiro intervalo a função varia de 0 até ½ enquanto no segundo intervalo a função varia de 1/2 até 0, não sendo em nenhum ponto dos intervalos negativa. Passado neste critério calculemos a integral da função neste intervalo.

$$\int_{-2}^{2} f(x)\,dx=\int_{-2}^{0}\frac{(2+x)}{4}dx+\int_{0}^{2}\frac{(2-x)}{4}dx=\int_{-2}^{0}\frac{2}{4}dx+\int_{-2}^{0}\frac{x}{4}dx+\int_{0}^{2}\frac{2}{4}dx-\int_{0}^{2}\frac{x}{4}dx \quad ,$$

ou

$$\int_{-2}^{2} f(x)\,dx=\frac{1}{2}x\Big|_{-2}^{0}+\frac{1}{4}\frac{x^2}{2}\Big|_{-2}^{0}+\frac{1}{2}x\Big|_{0}^{2}-\frac{1}{4}\frac{x^2}{2}\Big|_{0}^{2}=\frac{2}{2}-\frac{4}{8}+\frac{4}{4}-\frac{4}{8}=1 \quad ,$$

portanto é distribuição de probabilidade. Observemos que há outra maneira de demonstrarmos isto observando (basta fazer um desenho) que esta função é um triângulo de base 4 e altura 1/2, ou seja, tem área 1.
Calculemos agora a variância começando pela média, ou seja,

$$\mu=\int_{-2}^{2} x f(x)\,dx=\int_{-2}^{0} x\,\frac{(2+x)}{4}\,dx+\int_{0}^{2} x\,\frac{(2-x)}{4}\,dx=\int_{-2}^{0}\frac{2}{4}x\,dx+\int_{-2}^{0}\frac{x^2}{4}\,dx+\int_{0}^{2}\frac{2}{4}x\,dx-\int_{0}^{2}\frac{x^2}{4}\,dx \quad ,$$

ou ainda,

$$\mu=\frac{1}{2}\frac{x^2}{2}\Big|_{-2}^{0}+\frac{1}{4}\frac{x^3}{3}\Big|_{-2}^{0}+\frac{1}{2}\frac{x^2}{2}\Big|_{0}^{2}-\frac{1}{4}\frac{x^3}{3}\Big|_{0}^{2}=\frac{1}{4}-\frac{4}{8}+\frac{4}{4}-\frac{4}{8}=-1+\frac{2}{3}+1-\frac{2}{3}=0 \quad .$$

Novamente, este resultado também poderia ser obtido de forma mais rápida pois, olhando na figura correspondente a esta distribuição, veremos que a origem é a média da distribuição devido à simetria da distribuição em torno da origem.
Calculemos agora a integral

$$\int_{-2}^{2} x^2 f(x)\,dx=\int_{-2}^{0} x^2\frac{(2+x)}{4}\,dx+\int_{0}^{2} x^2\frac{(2-x)}{4}\,dx=\int_{-2}^{0}\frac{2}{4}x^2\,dx+\int_{-2}^{0}\frac{x^3}{4}\,dx+\int_{0}^{2}\frac{2}{4}x^2\,dx-\int_{0}^{2}\frac{x^3}{4}\,dx \quad .$$

Integrando obteremos

$$\int_{-2}^{2} x^2 f(x)\,dx=\frac{1}{2}\frac{x^3}{3}\Big|_{-2}^{0}+\frac{1}{4}\frac{x^4}{4}\Big|_{-2}^{0}+\frac{1}{2}\frac{x^3}{3}\Big|_{0}^{2}-\frac{1}{4}\frac{x^4}{4}\Big|_{0}^{2}=\frac{1}{2}\frac{8}{3}-\frac{1}{4}\left(\frac{16}{4}\right)+\frac{1}{2}\frac{8}{3}-\frac{1}{4}\frac{16}{4}=\frac{4}{3}-1+\frac{4}{3}-1=\frac{2}{3}$$

Como a média é zero, este valor é o valor da variância.

2 – Segunda questão (1,5 pontos)
Foi sorteada uma amostra de 20 postos de saúde da rede pública em uma determinada cidade e anotado o número de casos de dengue em cada uma deles no mês de março. Os resultados foram: 12, 9, 4, 7, 11, 3, 9, 1, 19, 3, 14, 8, 3, 21, 6, 9, 11, 13, 17, 7. Use os estimadores de média e variância apresentados abaixo para avaliar estes parâmetros estatísticos.

$$\mu_1=\frac{(valormínimo+valormáximo)}{2} \quad \text{E} \quad \mu_2=\bar{X} \qquad (0{,}5)$$

$$\sigma_1^2=\frac{(valormáximo)^2-(valormínimo)^2}{2} \quad \text{e} \quad \sigma_2^2=\frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2-n\,\bar{X}^2\right) . \qquad (0{,}5)$$

Use os dois valores obtidos para as médias para uso do estimador de variância que necessita desta informação.
Diga qual o estimador mais adequado para a média e para a variância. (0,5)
Resolução:
Para os estimadores da média teremos para o primeiro, por inspeção dos valores dados

$$\mu_1 = \frac{(valormínimo + valormáximo)}{2} = \frac{21+1}{2} = 11$$

e para o segundo

$$\mu_2 = \bar{X} = \frac{1}{20}\sum_1^{20} x_i = \frac{12+9+4+7+11+3+9+1+19+3+14+8+3+21+6+9+11+13+17+7}{20} = \frac{187}{20}$$

ou seja, $\mu_2 = 9{,}35$ **.**

Calculemos agora os estimadores para a variância.

$$\sigma_1^2 = \frac{(valormáximo)^2 - (valormínimo)^2}{2} = \frac{21^2 - 1^2}{2} = 220 \text{ ,}$$

sendo o segundo estimador dado por

$$\sigma_2^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n\,\bar{X}^2\right) \text{ .}$$

Repare que este estimador exige o uso do valor da média. Portanto, usaremos, como foi pedido, as duas estimativas das médias. Calculemos primeiro o somatório

$$\sum_1^{20} X_i^2 = 12^2 + 9^2 + 4^2 + 7^2 + 11^2 + 3^2 + 9^2 + 1^2 + 19^2 + 3^2 + 14^2 + 8^2 + 3^2 + 21^2 + 6^2 + 9^2 + 11^2 + 13^2 + 17^2 + 7^2$$

ou

$$\sum_1^{20} X_i^2 = 2327 \text{ .}$$

Pelo uso do segundo estimador teremos

$$\sigma_{22}^2 = \frac{1}{19}\left(\sum_{i=1}^{20} X_i^2 - 20\,\bar{X}_1^{\,2}\right) = \frac{1}{19}\left(2327 - 20 \times 9{,}35^2\right) = \frac{558{,}55}{19} = 30{,}45 \text{ .}$$

Usando o primeiro estimador teremos

$$\sigma_{21}^2 = \frac{1}{19}\left(\sum_{i=1}^{20} X_i^2 - 20\,\bar{X}_1^{\,2}\right) = \frac{1}{19}\left(2327 - 20 \times 11\right) = -\frac{93}{19} \text{ ,}$$

o que é absurdo pois a variância é sempre positiva. Claro, isto era uma possibilidade pois usamos um estimador para a média que era viciado. Este item do exercício tem por função chamar a atenção de inconsistências que podem surgir pelo uso descuidado dos estimadores.

Os estimadores de sub-índice 2 são os mais adequados pois são consistentes e não viciados.

3- Terceira questão (2,0 pontos)
Calcule as probabilidades abaixo.

a) P(X > 3,2) supondo que a distribuição é Uniforme no intervalo [-1, 8]. (0,5)
Resolução:
A probabilidade de uma distribuição Uniforme é dada por

$$P(a<X<b)=\frac{1}{B-A}\int_a^b dx \quad ,$$

onde [A, B] (B > A) é o intervalo onde está definida a distribuição. No caso teremos

$$P(X>3{,}2)=\frac{1}{8-(-1)}\int_{3{,}2}^{8} dx=\frac{1}{9}(8-3{,}2)=\frac{8}{15}\approx 0{,}5333$$

b) P(0,72 < X < 4,73) supondo que a distribuição segue o modelo Exponencial com $\alpha=0{,}75$.
Resolução:
Aqui a probabilidade será dada por

$$P(a<X<b)=\int_a^b \alpha e^{-\alpha x}\,dx=e^{-\alpha a}-e^{-\alpha b} \quad .$$

No caso específico teremos

$$P(0{,}72<X<4{,}73)=e^{-0{,}75\times 0{,}72}-e^{-0{,}75\times 4{,}73}\approx 0{,}5539 \quad .$$

c) P(-1,3 < X < 1,5) supondo que a distribuição seja

$$f(x)=\begin{matrix}(2+x)/4, & se\ -2\leq\ x\ \leq 0; \\ (2-x)/4, & se\ 0\leq\ x\ \leq 2.\end{matrix} \quad .$$

Resolução:
Aqui teremos uma definição mais complexa pelo fato da distribuição ser composta por duas funções e esta será

$$P(-1{,}3<X<1{,}5)=\int_{-1{,}3}^{1{,}5} f(x)\,dx=\int_{-1{,}3}^{0}\frac{(2+x)}{4}dx+\int_{0}^{1{,}5}\frac{(2-x)}{4}dx=\int_{-1{,}3}^{0}\frac{2}{4}dx+\int_{-1{,}3}^{0}\frac{x}{4}dx+\int_{0}^{1{,}5}\frac{2}{4}dx-\int_{0}^{1{,}5}\frac{x}{4}dx$$

ou seja,

$$P(-1{,}3<X<1{,}5)=\frac{1}{2}x|_{-1{,}3}^{0}+\frac{1}{4}\frac{x^2}{2}|_{-1{,}3}^{0}+\frac{1}{2}x|_{0}^{1{,}5}-\frac{1}{4}\frac{x^{1{,}5}}{2}|_{0}^{2}=\frac{13}{20}-\frac{169}{800}+\frac{3}{4}-\frac{9}{32}=\frac{363}{400}=0{,}9075$$

d) P(0,73<X<2,18), distribuição Normal, dado média igual a 0,9 e variância 4,41.
Resolução:

Agora teremos a probabilidade dada por

$$\boldsymbol{P}(a \le X \le b) = P\left(\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}\right) \quad ,$$

que no nosso caso dará em

$$P(0{,}73 \le X \le 2{,}18) = P\left(\frac{0{,}73-0{,}9}{\sqrt{4{,}41}} \le Z \le \frac{2{,}18-0{,}9}{\sqrt{4{,}41}}\right) = P\left(\frac{-0{,}17}{2{,}1} \le Z \le \frac{1{,}28}{2{,}1}\right) \approx P(-0{,}0809 < Z < 0{,}6095)$$

.

Usando as propriedades da distribuição Normal teremos

$$P(-0{,}0809 \le X \le 0{,}6095) \approx P(Z < 0{,}08) + P(Z < 0{,}61) = 0{,}0318 + 0{,}2290 = 0{,}2608 \quad .$$

4- Quarta questão (2,0 pontos)

Numa experiência de campo levantou-se 10 amostras com os valores abaixo em unidades arbitrárias.

L	8,1	12,5	12,5	18,2	23,4	18,8	14,9	15,4	19,2	18,3

Usando estimadores não viciados, calcule o intervalo de confiança para a valor médio da amostra com um coeficiente de confiança de 75%.

Resolução:

O intervalo de confiança é dado por

$$IC(\mu,\gamma) = \left[\bar{X} - z_{\gamma/2}\frac{\sigma}{\sqrt{n}} \,;\, \bar{X} + z_{\gamma/2}\frac{\sigma}{\sqrt{n}}\right] .$$

Será, portanto, determinar estimativas para a variância. Usaremos, pois, estimadores consistentes e não viciados para obter este valor. Calculemos a média

$$\bar{X} = \frac{1}{10}\sum_{1}^{10} x_i = 8{,}1+12{,}5+12{,}5+18{,}2+23{,}4+18{,}8+14{,}9+15{,}4+19{,}2+\frac{18{,}3}{10} = 16{,}13 \quad ,$$

e então calculemos o somatório

$$\sum_{1}^{10} x_i^2 = 8{,}1^2+12{,}5^2+12{,}5^2+18{,}2^2+23{,}4^2+18{,}8^2+14{,}9^2+15{,}4^2+19{,}2^2+18{,}3^2 = 2773{,}05 \ .$$

Daqui podemos calcular a variância como

$$\sigma^2 = \frac{1}{9}\left(\sum_{i=1}^{10} X_i^2 - n\,\bar{X}^2\right) = \frac{1}{9}\left(2773{,}05 - 10 \times 16{,}13^2\right) = \frac{171{,}281}{9} \approx 19{,}0312 \ ,$$

logo $\sigma = \sqrt{19{,}0312} \approx 4{,}3624$.

Com estes valores podemos calcular o intervalo de confiança

$$IC(\mu,0,75)=\left[16,13-z_{0,375}\frac{4,3624}{\sqrt{10}};16,13+z_{0,375}\frac{4,3624}{\sqrt{10}}\right].$$

Usando a tabela de distribuição Normal obtemos que $z_{0,375}=1,15$ **. Daí**

$$IC(\mu,0,75)=\left[16,13-1,15\frac{4,3624}{3,1623};16,13+1,15\frac{4,3624}{3,1623}\right]=[14,5435,17,7164]\approx[14,54,17,72]$$

.

5 – Quinta questão (1,0 pontos)

O comprimento de uma peça produzida numa fábrica tem média 30 cm e variância 4 cm^2 sendo que a amostra estudada era de 30 peças. Esta produção foi modelada pela distribuição Normal. Calcule as seguintes probabilidades:

a) De uma peça ter comprimento maior que 31 cm; (0,5 ponto)

Resolução:

Usaremos, já com os parâmetros dados, a fórmula

$$\boldsymbol{P}(X>31)=P\left(Z>\frac{31-30}{2}\right)=P\left(Z>\frac{1}{2}\right)=0,5-P\left(Z<\frac{1}{2}\right)=0,5-0,1915=0,3085 \quad .$$

b) De uma peça ter comprimento entre 29 cm e 32 cm; (0,5 ponto)

Resolução:

$$\boldsymbol{P}(29<X<32)=P\left(\frac{29-30}{2}<Z<\frac{32-30}{2}\right)=P\left(-\frac{1}{2}<Z<1\right)=P\left(Z<\frac{1}{2}\right)+P(Z<1)$$

e daí

$$P(29<X<32)=0,1915+0,3413=0,5328 \quad .$$

6 - Sexta questão (1,0 pontos)
Durante um trabalho na zona rural foi detectada a contaminação do lençol freático por nitratos, devido ao uso inadequado de adubos nitrogenados. Foram extraídas amostras de água de 20 poços com média amostral de 12,4 mg por litro. Testes similares indicam uma variância de 6,4. Determine qual é o intervalo de confiança para a média com coeficiente de confiança de 75%. E com coeficiente de confiança de 95%?

Resolução:

Novamente partiremos da fórmula

$$IC(\mu,\gamma)=\left[\bar{X}-z_{\gamma/2}\frac{\sigma}{\sqrt{n}}\,;\bar{X}+z_{\gamma/2}\frac{\sigma}{\sqrt{n}}\right],$$

que pode ser escrita com parte dos dados do problema como

$$IC(\mu,\gamma)=\left[12{,}4-z_{\gamma/2}\frac{\sqrt{6{,}4}}{\sqrt{20}}\,;12{,}4+z_{\gamma/2}\frac{\sqrt{6{,}4}}{\sqrt{20}}\right]\approx\left[12{,}4-z_{\gamma/2}0{,}5656\,;12{,}4+z_{\gamma/2}0{,}5656\right].$$

Pelo uso da tabela da distribuição Normal obtemos que $z_{0,75/2}=z_{0,375}=1{,}15$ **e** $z_{0,95/2}=z_{0,475}=1{,}96$ **.**

Substituindo em cada caso teremos

$$IC(\mu,0{,}75)=[12{,}4-1{,}15\times0{,}5656\,;12{,}4+1{,}15\times0{,}5656]\approx[11{,}75\,;13{,}05]$$ **e**

$$IC(\mu,0{,}95)=[12{,}4-1{,}96\times0{,}5656\,;12{,}4+1{,}96\times0{,}5656]\approx[11{,}29\,;13{,}51].$$

Como já era esperado, ao aumentarmos o coeficiente de confiança aumentamos o intervalo de confiança.