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

Curso de Tecnologia em Sistemas de Computação
Disciplina Probabilidade e Estatística

Gabarito da
AP2 2° semestre de 2006.

Professores: Otton Teixeira da Silveira Filho e Regina Célia P. Leal Toledo

1 - Primeira questão

Dado o gráfico abaixo (onde não houver indicação a função vale zero):

a) Demonstre que f(x) é uma densidade;

f(x) é sempre positiva e a área sob a curva vale 1. A área pode ser obtida pelas fórmulas das áreas do retângulo e do trapézio ou ainda por integração.

b) Escreva a expressão da função;

$$\frac{1}{5};x\in[0,1)$$
$$\frac{1}{15}(2x+1);x\in[1,2)$$
$$\frac{1}{15}(-x+7);x\in[2,4]$$

c) Calcule o valor médio;

Pela definição de valor médio para funções contínuas

$$\mu=\int_{-\infty}^{\infty} x f(x)dx$$

teremos:

$$\mu=\int_0^1 \frac{x}{5}dx+\int_1^2 \frac{x}{15}(2x+1)dx+\int_2^4 \frac{x}{15}(-x+7)dx=2,066667$$

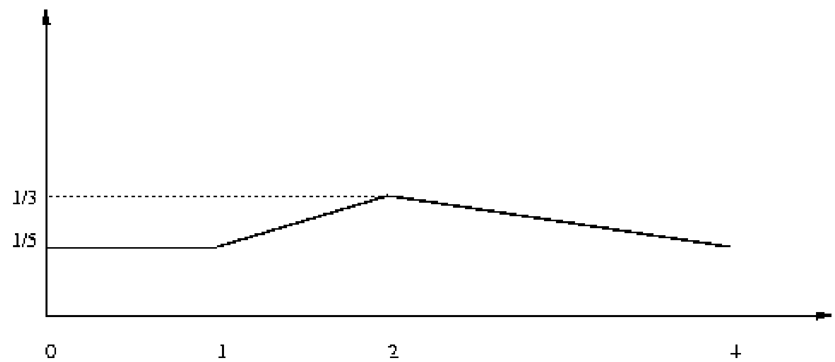
d) Calcule a variância;

A definição de variância é

$$\sigma^2 = \int_{-\infty}^{\infty} x^2 f(x) dx$$

$$\sigma^2 = \int_0^1 \frac{x^2}{5} dx + \int_1^2 \frac{x^2}{15}(2x+1) dx + \int_2^4 \frac{x^2}{15}(-x+7) dx = 5,433333$$

e) Calcule a moda.



Por observação direta da figura temos que a moda (valor máximo da função densidade) é 1/3.

2 - Segunda questão (2,0 pontos)

Uma função densidade de probabilidade varia da forma f(x) = Cx no intervalo [0,2] valendo zero fora do intervalo.

a) Detemine a constante C; (0,5 ponto)

A função acima é uma função de probabilidade se C por positiva e não nula e tal que a área abaixo da função seja igual a um. A área é a de um triângulo e esta deve ser igual a 1 para ser densidade de probabilidade. Por integração temos

$$\int_0^2 Cx\, dx = C \frac{x^2}{2} \Big|_0^2 = 2C$$

ou seja, C deve ser igual a 1/2

b) Calcule:

Todas as questões abaixo podem ser obtidas por integração ou pelo uso da fórmula da área do trapézio e pelas propriedades de complementaridade.

I)$P(X > 1/3)$ = 35/36

II)$P(X < 1/3)$ = 1/36

III)$P(1/3 < X < 2)$ = 35/36

3 - Terceira questão (2,0 pontos)

Se investiga numa empresa de informática a possibilidade de mudança da linguagem na qual são elaboradas as aplicações desenvolvidas. O fator avaliado é a velocidade de implementação nas linguagens. Uma equipe de programadores, experientes na linguagem usada atualmente (linguagem 1) e numa proposta para novo uso (linguagem 2), desenvolveram programas a partir do mesmo algorítmo. Abaixo vai uma tabela com os resultados de tempos obtidos para o desenvolvimento de programas baseados no algorítmo.

linguagem 1	linguagem 2
17	18
16	14
21	19
14	11
18	23
24	21
16	10
14	13
21	19
23	24
13	15
18	20
22	18
14	16
20	21

a) Calcule um intervalo de confiança de 95% para a diferença das médias no tempo de programação;

Examinando as amostras, temos que a média aritmética de tempo de elaboração do algoritmo na linguagem 1 é aproximadamente $\mu_1 = 18,07$.
Para obtermos uma estimativa para variância, usaremos a variância amostral dada por

$$S^2 = \frac{1}{n-1} \sum_1^n X_i^2 - n \bar{X}^2 b = \frac{1}{14} \sum_1^{15} X_i^2 - 15 \bar{X}^2$$

que para a linguagem 1 nos dá:

$$S^2 = 12,92 \Rightarrow S \approx 3,59$$

Vamos supor que o tamanho da amostra é grande o suficiente para que possamos usar o Teorema do Limite Central. Daí podemos usar para a distribuição

$$N(18,07 ; 3,59)$$

Sendo o intervalo de confiança dado por

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

para o valor de confiança igual a $\gamma = 0,95$ retiramos da tabela da distribuição Normal o valor de $z_{\gamma/2} = 1,96$. Daí teremos

$$IC(\mu, 0,95) = \left[18,07 - 1,96 \frac{3,59}{\sqrt{15}} ; 18,07 + 1,96 \frac{3,59}{\sqrt{15}} \right] = [16,25 ; 19,88]$$

b) Por estes dados podemos concluir que a linguagem deve ser mudada?

Observe que a média para a linguagem 2 é $\mu_1 = 17,47$ que está dentro do intervalo de confiança. Isto indica que dentro do valor de confiança requerido, a mudança de linguagem não tem vantagens.

4 - Quarta questão (1,5 pontos)

Uma firma que vende pela Internet afirmou que 40% dos seus clientes tem escolaridade de nível superior. Um concorrente tinha uma pesquisa nas qual sabia quais dos clientes as duas empresas tinham em comum. Baseados nestes clientes, colheu uma amostra de 50 leitores para verificar a afirmação do concorrente.

a) Escreva este problema como um teste de hipóteses e interprete os erros dos tipos I e II;

H_0 : Não existem 40 % de clientes com escolaridade de nível superior

H_a : Existem tais clientes

b) Determine a região crítica com um nível de significância de 10%;

Usaremos com a proporção amostral, já que estamos trabalhando com a condição dos clientes terem ou não nível de escolaridade superior. Assim, obtemos as estimativas para a média e a variância iguais a

$$\hat{p}=\frac{\text{número de indivíduos na amostra com dada característica}}{n}=0,4$$

$$Var(\hat{p})=\frac{p(1-p)}{n}=0,40\frac{(1-0,40)}{50}=0,0048$$

$$\alpha=P(\text{erro tipo I})=P\left(\bar{X}<x_c|\mu=0,40\right)=P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}<\frac{x_c-0,40}{0,069/\sqrt{50}}\right)$$

$$\alpha=P\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}<\frac{x_c-0,40}{0,01}\right)=P(Z<z_c)$$

Para obtemos x_c da seguinte equação

$$z_c=\frac{x_c-0,40}{0,01}$$

Para $\alpha=0,10$ temos

$$0,10=P(Z<z_c)\Rightarrow z_c=1,28$$

Daí

$$x_c = 0{,}0128 + 0{,}40 = 0{,}4128$$

A região crítica é, portanto,

$$RC = \{x \in \mathbb{R} : X < 0{,}4128\}$$

c) Baseado no calculado acima, qual a probabilidade de 21 clientes responderem que tem escolaridade que não seja de nível superior?

<u>Questão anulada: os itens a e b ficam com o total da quarta questão.</u>

5- Quinta questão (2,0 pontos)

Numa eleição concorriam dois candidatos (A e B). Um deles, o candidato A, obteve 60% dos votos. Numa seção eleitorial, julgada representativa do eleitorado, qual a probabilidade de do candidato A ter conseguido entre 550 e 730 votos de 1000 votos? (1,0 ponto) Qual a probabilidade do candidato B ter conseguido o mesmo número de votos? (1,0 ponto)

Partindo da suposição de que a amostra é grande o suficiente para usarmos o Teorema Central do Limite, trabalharemos com o modelo Normal. Pela natureza do problema, usaremos a proporção amostral

$$\hat{p} = \frac{\text{número de indivíduos na amostra com dada característica}}{n}$$

e o estimador para a variância

$$Var(\hat{p}) = \frac{p(1-p)}{n}$$

A distribuição pelo Teorema Central do Limite será

$$\hat{p} \sim N\left(0{,}6 ; 0{,}6\frac{(1-0.6)}{1000}\right) = N(0{,}6 ; 0{,}00024)$$

Observe que pelo valor da variância o resultado final se anuncia.
A probabilidade do candidato A ter obtido entre 550 e 730 votos em 1000 votos será

$$P(0,55<X<0,73)\approx P\left(\frac{0,55-0,6}{\sqrt{\frac{0,4(1-0,6)}{1000}}}<\frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}}<\frac{0,73-0,6}{\sqrt{\frac{0,6(1-0,6)}{1000}}}\right)$$

$$P(0,55<X<0,73)\approx P\left(-3,2274<\frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}}<8,3914\right)$$

$$P(0,55<X<0,73)\approx P\left(\frac{-3,2274-0,6}{0,01549}<Z<\frac{8,3914-0,6}{0,01549}\right)\text{¿¿}$$

$$P(0,55<X<0,73)\approx P(-247,05<Z<502,93)\approx 1$$

ou seja, é quase que certo que o número de votos estará nesta faixa.

b) No caso do candidato B, temos pela complementaridade que a probabilidade de ter a votação, dentro da faixa de votação proposta, é virtualmente nula.