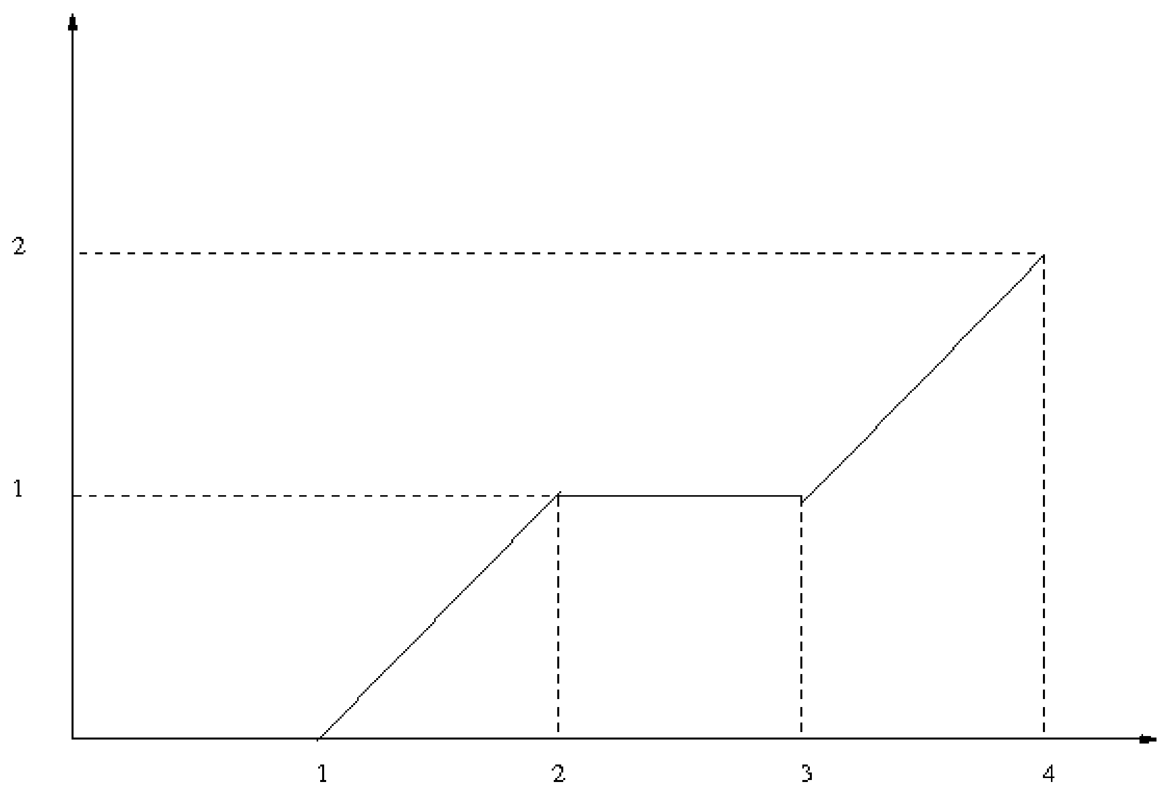
Fundação CECIERJ - Vice Presidência de Educação Superior a Distância

Curso de Tecnologia em Sistemas de Computação
Disciplina Probabilidade e Estatística
AD2 2° semestre de 2011
GABARITO

1 - Primeira questão (1,5 pontos)

Abaixo é apresentado o gráfico de uma distribuição que deve ser normalizada, ou seja, multiplicada por uma constante C de forma que a sua integral seja igual a 1. Para valores de x maiores que quatro ou menores que 1 a função é nula.



a) Calcule a constante de normalização; (0,5)

RESOLUÇÃO:

Para este item existem pelo menos duas maneiras de resolver.

A primeira é perceber que a área total é a soma da área de um triângulo, um retângulo e um trapézio. O triângulo vai de x = 1 até x = 2 com altura 1 com área A_{trig}. O retângulo vai de 2 até 3 com altura 1 e área A_{ret}. O trapézio tem base maior

igual a 2 e base menor igual a 1, altura 1 (olhe a figura de lado) e com área A_{trap}. Assim, a área total, baseando no valores do gráfico será

$$A=A_{trig}+A_{ret}+A_{trap}=\frac{1\times 1}{2}+1\times 1+\frac{1+2}{2}\times 1=\frac{1}{2}+1+\frac{3}{2}=3$$

portanto a constante de normalização é $1/3$ **.**

Na segunda maneira se determina as equações de cada segmento contínuo da distribuição, o que nos será útil no item b.

A equação da reta é dada por $y=ax+b$ **. Assim a reta (lado do triângulo) que passa pelos pontos (1,0) e (2,1) será dada pela solução do sistema de equações**

$$\begin{aligned} a\times 2+b&=1 \\ a\times 1+b&=0 \end{aligned} \ .$$

Resolvendo o sistema teremos $y=x-1$ **.**

Para a segunda reta (lado do outro triângulo) temos que a reta passa pelos pontos (3,1) e (4,2). Novamente criando o sistema, teremos para este pontos

$$\begin{aligned} a\times 4+b&=2 \\ a\times 3+b&=1 \end{aligned}$$

Que resolvido nos dará para a equação desta reta $y=x-2$

Assim teremos

$$y=x-1\,;x\in[1,2) \qquad y=1\,;x\in[2,3) \qquad y=x-2\,;x\in[3,4)$$

Integrando teremos

$$\int_{-\infty}^{\infty} F(x)dx=\int_{1}^{4} f(x)dx=\int_{1}^{2}(x-1)\,dx+\int_{2}^{3} 1\,dx+\int_{3}^{4}(x-2)\,dx \ .$$

ou

$$\int_{-\infty}^{\infty} F(x)dx=(\frac{x^2}{2}-x)|_1^2+x|_2^3+(\frac{x^2}{2}-2x)|_3^4=\frac{1}{2}+1+\frac{3}{2}=3$$

Assim a constante de normalização é $\frac{1}{3}$ **.**

b) Feito isto calcule a média da distribuição obtida; (0,5)

RESOLUÇÃO:

Partindo da definição de média de distribuições contínuas teremos a resolução do problema, ou seja,

$$\int_{-\infty}^{\infty} x\,F(x)\,dx=\frac{1}{3}\int_{0}^{4} x\,f(x)\,dx=\frac{1}{3}\left(\int_{1}^{2} x(x-1)\,dx+\int_{2}^{3} x\,dx+\int_{3}^{4} x(x-2)\,dx\right)$$

ou ainda

$$\int_{-\infty}^{\infty} x\,F(x)\,dx=\frac{1}{3}\left[\left(\frac{x^3}{3}-\frac{x^2}{2}\right)\Big|_1^2+\frac{x^2}{2}\Big|_2^3+\left(\frac{x^3}{3}-2\frac{x^2}{2}\right)\Big|_3^4\right]=\frac{1}{3}\left[\frac{5}{6}+\frac{5}{2}+\frac{16}{3}\right]=\frac{1}{3}\,\frac{26}{3}=\frac{26}{9}=2{,}888...$$

c) Calcule a variância da distribuição obtida. (0,5).

RESOLUÇÃO:
Aqui partiremos da definição de variância para distribuições contínuas

$$\sigma^2=\int_{-\infty}^{\infty} x^2 F(x)\,dx-\mu^2$$

Calculemos a integral

$$\int_{-\infty}^{\infty} x^2 F(x)dx=\frac{1}{3}\int_{0}^{4} f(x)\,dx=\frac{1}{3}\left(\int_{1}^{2} x^2(x-1)\,dx+\int_{2}^{3} x^2\,dx+\int_{3}^{4} x^2(x-2)\,dx\right)$$

$$\int_{-\infty}^{\infty} x^2 F(x)dx=\frac{1}{3}\left[\left(\frac{x^4}{4}-\frac{x^3}{3}\right)\Big|_1^2+\frac{x^3}{3}\Big|_2^3+\left(\frac{x^4}{4}-2\frac{x^3}{3}\right)\Big|_3^4\right]=\frac{1}{3}\left[\frac{17}{12}+\frac{19}{3}+\frac{229}{12}\right]=\frac{1}{3}\,\frac{161}{6}=\frac{161}{18}$$

Finalmente a variância será

$$\sigma^2=\int_{-\infty}^{\infty} x^2 F(x)\,dx-\mu^2=\frac{161}{18}-\left(\frac{26}{9}\right)^2=\frac{97}{162}=0{,}5987$$

2 – Segunda questão (2,0 pontos)
Calcule as probabilidades abaixo na suposição que a distribuição de probabilidades segue a distribuição Normal.

RESOLUÇÃO:
Para uma distribuição Normal a probabilidade é dada por

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

ou de outro modo

$$P(X>a)=P\left(\frac{X-\mu}{\sigma}>\frac{a-\mu}{\sigma}\right)=P\left(Z>\frac{a-\mu}{\sigma}\right)\ .$$

a) **P(X> 3,3)**, dado média igual a 3,2 e variância 4,4 (0,5)

$$P(X>3{,}3)=P\left(\frac{X-\mu}{\sigma}>\frac{3{,}3-3{,}2}{\sqrt{4{,}4}}\right)=P(Z>0{,}047)\approx P(Z>0{,}05)=0{,}5-0{,}0199=0{,}4801$$

b) **P(2,2<X<3,7)**, dado média igual a 2,8 e variância 6,2 (0,5)

$$P(2{,}2\leq X\leq 3{,}7)=P\left(\frac{2{,}2-2{,}8}{\sqrt{6{,}2}}\leq Z\leq\frac{3{,}7-2{,}8}{\sqrt{6{,}2}}\right)=P(-0{,}241\leq Z\leq 0{,}361)$$

$$P(-0{,}24\leq Z\leq 0{,}36)=P(Z>0{,}24)+P(Z>0{,}36)=0{,}0948+0{,}1406=0{,}2354$$

c) **P(0,7<X<2,1)**, dado média igual a 1,6 e variância 1,1 (0,5)

$$P(0{,}7\leq X\leq 2{,}1)=P\left(\frac{0{,}7-1{,}6}{\sqrt{1{,}1}}\leq Z\leq\frac{2{,}1-1{,}6}{\sqrt{1{,}1}}\right)=P(-0{,}8581\leq Z\leq 0{,}4767)$$

$$P(-0{,}86\leq Z\leq 0{,}48)=P(Z>0{,}86)+P(Z>0{,}48)=0{,}3051+0{,}1844=0{,}4895$$

d) **P(8,7<X<12,1)**, dado média igual a 8,6 e variância 4,6 (0,5)

$$P(8{,}7\leq X\leq 12{,}1)=P\left(\frac{8{,}7-8{,}6}{\sqrt{4{,}6}}\leq Z\leq\frac{12{,}1-8{,}6}{\sqrt{4{,}6}}\right)=P(0{,}0466\leq Z\leq 1{,}6318)$$

$$P(0{,}05\leq Z\leq 1{,}63)=P(Z<1{,}63)-P(0{,}05)=0{,}4484-0{,}0199=0{,}4285$$

3- Terceira questão (1,5 pontos)
Calcule as probabilidades abaixo.

a) P(X > 4) supondo que a distribuição é uniforme no intervalo [1,6]. (0,5)
RESOLUÇÃO:
Esta distribuição é dada por

$$\frac{1}{b-a},\ a\leq x\leq b\,;\ 0\ \textit{fora deste intervalo}$$

No nosso caso a = 1 e b = 6 e teremos a distribuição

$$\frac{1}{6-1}=\frac{1}{5},\ 1\leq x\leq 6\quad.$$

Assim temos que a probabilidade pedida será

$$P(X>4)=\int_4^6\frac{1}{5}\,dx=\frac{1}{5}x\Big|_4^6=\frac{6-4}{5}=\frac{2}{5}=0{,}4$$

ou, de outra forma, a área do retângulo de altura 1/5 e largura 2.

b) P(1 < X < 2) supondo que a distribuição segue o modelo Exponencial com $\alpha=1,2$. (0,5)
RESOLUÇÃO:
Para a distribuição Exponencial de parâmetro α a probabilidade é dada por

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

No caso presente ficaremos então com

$$P(1<X<2)=e^{-1,2\times1}-e^{-1,2\times2}\approx0,3012-0,0907=0,2105 \quad .$$

c) P(1,5 < X < 2,5) supondo que a distribuição é a encontrada no item **a** da primeira questão. (0,5)
RESOLUÇÃO:
A probabilidade será dada por

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

ou ainda

$$P(1,5<X<2,5)=\frac{1}{3}\left[\left(\frac{x^2}{2}-x\right)\Big|_{1,5}^{2}+x\Big|_{2}^{2,5}\right]=\frac{1}{3}(0,375+0,5)=0,2916 \quad .$$

4- Quarta questão (1,5 pontos)

Foram colhidas amostras de sangue de pessoas que viviam numa zona rural para uma investigação sobre contaminação por um determinado defensivo agrícola. Numa avaliação preliminar foram colhidas amostras de 15 pessoas e os valores medidos estão apresentados abaixo em unidades arbitrárias (UA):

P	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
UA	4,33	6,31	7,22	8,01	5,12	6,11	4,92	5,61	6,77	8,31	9,02	6,15	5,08	6,72	7,56

Faça as hipóteses cabíveis à situação e calcule o intervalo de confiança para a média real com um coeficiente de confiança de 95%.

RESOLUÇÃO:
O intervalo de confiança é dado por

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

Portanto são necessárias a média e a variância, no caso amostrais, desta massa de dados. Adotaremos dois estimadores não viciados.

A média amostral é $\bar{X}=\frac{\sum_{1}^{n} x_i}{n}=\frac{97,24}{15}\approx 6,4826$ **enquanto a estimativa para a**

variância amostral é dada por

$$\sigma^2=\frac{1}{n-1}\left(\sum_{1}^{n} X_i^2-n\bar{X}^2\right)=\frac{1}{14}(656,2688-15\times 42,0250)$$

ou

$$\sigma^2=\frac{1}{14}(656,2688-630,3745)\approx 1,8495 \text{ ,}$$

ou ainda, $\sigma\approx\sqrt{1,8495}\approx 1,36$ **.**

Para 95% de coeficiente ($\gamma=0,95$ **) de confiança teremos** $z_{\gamma/2}=z_{0,475}=1,96$ **, que é achado lendo a tabela da função Normal. Assim, o intervalo de confiança para a média será**

$$IC(\mu, 0,95)=\left[6,4826-\frac{1,96\times 1,36}{\sqrt{15}} ; 6,4826+\frac{1,96\times 1,36}{\sqrt{15}}\right]=[5,79 ; 7,17]$$

5 – Quinta questão (1,5 pontos)
Foi feita uma pesquisa sobre escolaridade em adultos numa determinada cidade. Se estabeleceu que a média da escolaridade em anos é de 7,3 anos e estimou-se que a variância é 4,9 anos^2. Qual a probabilidade de encontrarmos pessoas de escolaridade no intervalo [7,6; 9] anos numa amostra de 10 pessoas?

RESOLUÇÃO:
Aqui vamos supor que a distribuição seja Normal. Como foram 10 pessoas teremos a distribuição $N(7,3 ; 4,9/10)=N(7,3 ; 0,49)$ **e assim** $\sigma=\sqrt{0,49}=0,7$

$$P(7,6\leq X\leq 9)=P\left(\frac{7,6-7,3}{0,7}\leq Z\leq\frac{9-7,3}{0,7}\right)=P(0,428\leq Z\leq 2,428)\approx P(0,43\leq Z\leq 2,43)$$

daí

$$P(0,43\leq Z)-P(Z<2,43)=(0,5-0,1664)-(0,5-0,4925)=0,3336-0,0075=0,3261 \text{ .}$$

6 - Sexta questão (2,0 pontos)
Numa cidade se avaliava a aplicação de uma lei anti-fumo. A medida era polêmica nesta cidade e políticos avaliaram que somente se pelo menos 70% da população apoiasse é que haveria a apresentação do projeto para a implementação da lei. Para evitar desgaste político, tomou-se uma amostra aleatória de 200 pessoas e verificou-se que 136 aprovavam a medida e as restantes 64 não concordavam.
a) Estabeleça a hipótese nula e a alternativa;

RESOLUÇÃO:
H_0 : $\mu = 0{,}70$ **;**
H_a : $\mu > 0{,}70$ **.**

b) Com um nível de significância de 0,05 estabeleça se o projeto será apresentado.
RESOLUÇÃO:
Temos o valor do nível de significância mas não temos os valores de variância. Para este cálculo usaremos a frequência amostral dada por

$$\hat{p} = \frac{\text{número de indiviíduos com uma dada característica}}{\text{número total de indivíduos}} \quad \mathbf{e} \quad Var = \frac{p(1-p)}{n} \quad .$$

No presente caso o valor para p é 0,70 e daí

$$Var = \frac{p(1-p)}{n} = \frac{0{,}7 \times 0{,}3}{200} = 0{,}00105 \approx 0{,}001$$

Para determinar o valor crítico temos que

$$0{,}05 = P(\bar{X} < x_c | \mu = 0{,}7) = P\left(\frac{\bar{X} - \mu}{var} < \frac{x_c - 0{,}7}{0{,}001}\right) = P(Z < z_c) \quad .$$

Temos que

$$z_c = \frac{x_c - 0{,}7}{0{,}001} \Rightarrow x_c = 0{,}7 + z_c 0{,}001 \quad .$$

No entanto, para o valor de α **dado, temos que verificar quanto vale** z_c **. Para isto vamos à tabela da distribuição Normal, lendo a partir dos valores de probabilidade , e procuremos pelo complemento da probabilidade 0,05, ou seja, 0,45. O valor mais próximo na tabela é 0,4495 que corresponde ao valor 1,64. No entanto, como pegamos o complemento, o valor que usaremos será -1,64.**
Assim teremos

$$x_c = 0{,}7 - 1{,}64 \times 0{,}001 \approx 0{,}698$$

o que marginalmente dentro da Região de Rejeição. Portanto, o projeto não deverá ser apresentado...

Atenção: Não haverá formulário na segunda avaliação presencial.