Probabilidade

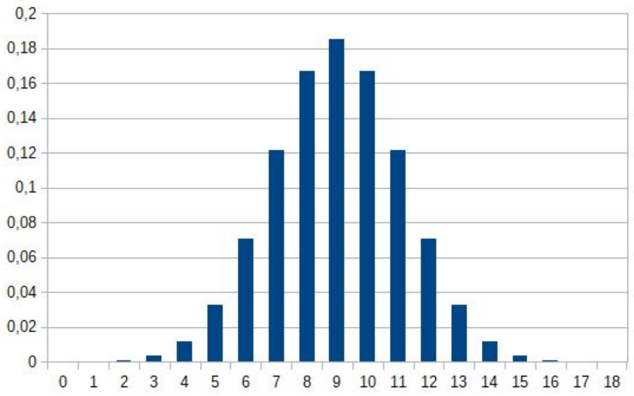
Distribuição contínua II

Prof. Dr. Tetsu Sakamoto
Instituto Metrópole Digital - UFRN
Sala A224, ramal 182
Email: tetsu@imd.ufrn.br

Slides e notebook em:

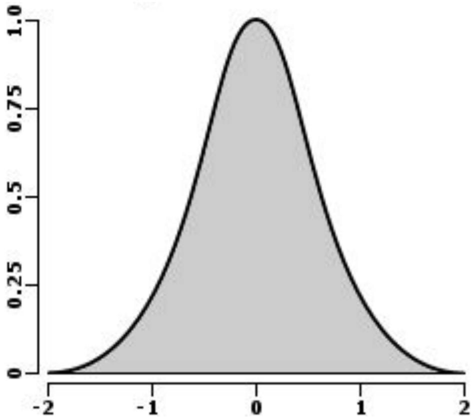
github.com/tetsufmbio/IMD0033/

Na aula passada



Função massa de probabilidade P(x)

- $P(x) \geq 0$;
- $\sum P(x_i) = 1$;

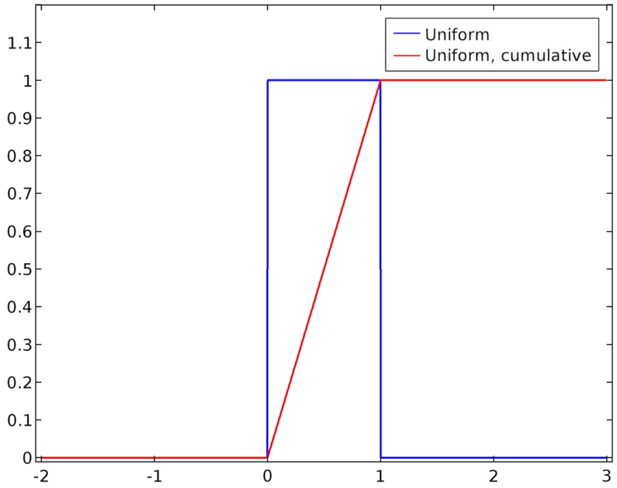


Função densidade de probabilidade f(x)

- $f(x) \geq 0$;
- Área sob a curva = 1;

Função de distribuição acumulada

$$F(x) = P(X < x) = \int_{-\infty}^{x} f(x)dx$$



$$f(x) \begin{cases} 1 \text{ se } 0 < x < 1 \\ 0 \end{cases}$$

F(x) = x

$$f(x) = F(x)'$$

Exemplo

Suponha que o tempo de vida de um componente eletrônico em meses seja uma variável aleatória contínua com $f(x) = 10/x^2$, $x > 10$.

1. Determine $P(X = 20)$;
2. Encontre a função de distribuição acumulada;
3. Determine $P(X < 20)$;
4. Determine a probabilidade de entre 6 desses componentes, 2 deles funcionarem por mais que 20 meses;

Exemplo

$f(x) = 10/x^2$, $x > 10$

Encontre a função de distribuição acumulada;

Exemplo

$f(x) = 10/x^2$, $x > 10$

Determine $P(X < 20)$;

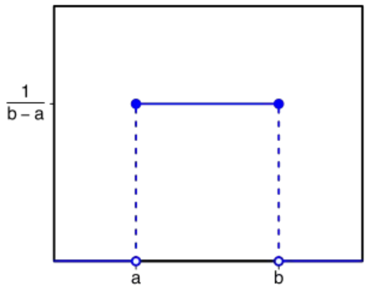
Exemplo

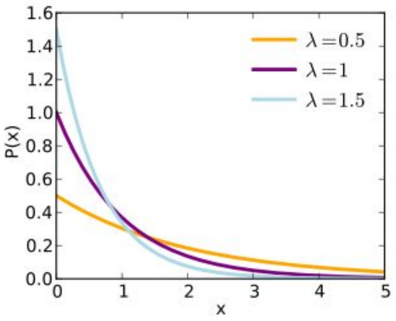
$f(x) = 10/x^2, x > 10$

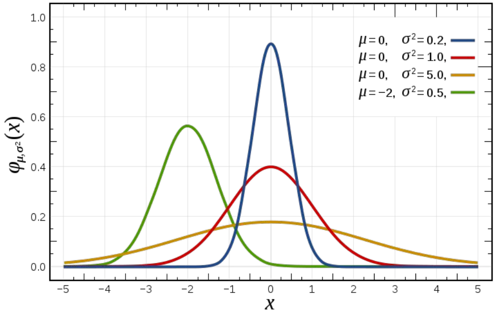
Determine a probabilidade de entre 6 desses componentes, 2 deles funcionarem por mais de 20 meses;

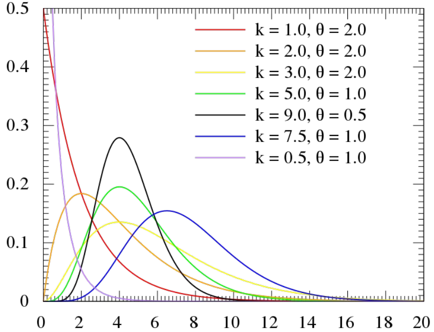
Famílias de distribuição de variáveis aleatórias contínuas

- Uniforme;
- Normal (Gaussiana);
- Exponencial;
- Gama;
- etc...





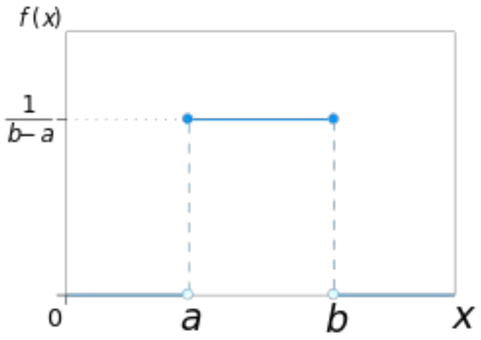


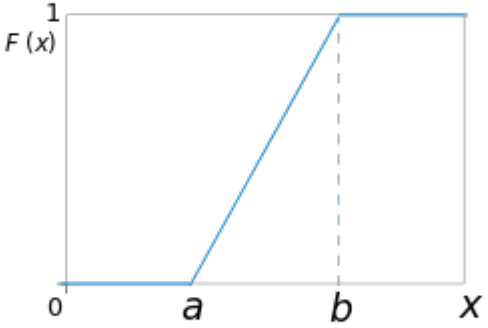


Distribuição Uniforme

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } x \in [a,b] \\ 0 & \text{otherwise} \end{cases}$$

$$F(x) = \begin{cases} 0 & \text{for } x < a \\ \frac{x-a}{b-a} & \text{for } x \in [a,b) \\ 1 & \text{for } x \geq b \end{cases}$$





$$E(X) = \frac{1}{2}(a+b)$$

$$V(X) = \frac{1}{12}(b-a)^2$$

Exemplo

Um trabalhador pode chegar no seu local de trabalho em qualquer momento entre 6 e 7 da manhã com a mesma probabilidade.

a) Esquematize o gráfico da função densidade de probabilidade da variável que mede o horário de chegada desse trabalhador.
b) Esquematize um gráfico da função acumulada da distribuição.
c) Calcule a probabilidade dele chegar antes de 6:15 e depois de 6:30.
d) Qual o horário esperado da sua chegada?

Distribuição exponencial

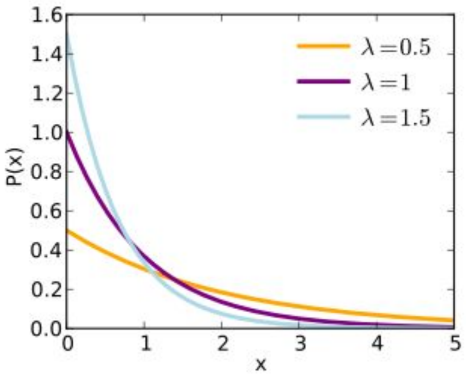
Análogo contínuo da distribuição geométrica;

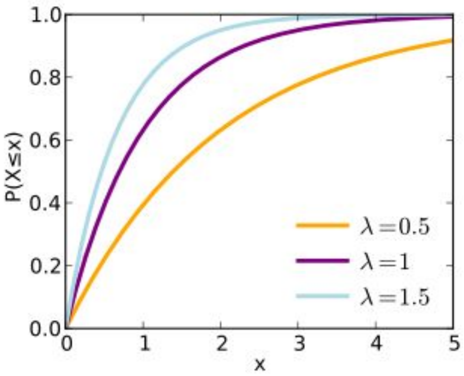
$$f(x) = \lambda e^{-\lambda x}, \lambda > 0$$

$$F(x) = 1 - e^{-\lambda x}$$

$$E(X) = \frac{1}{\lambda}$$

$$V(X) = \frac{1}{\lambda^2}$$





Exemplo

Considere X como sendo uma variável aleatória que expressa o tempo de espera (em minutos) para que um cliente seja atendido em uma padaria. É conhecido que X segue uma distribuição exponencial com a média de tempo igual a 4 minutos.

a) Esquematize a função densidade de probabilidade de X;
b) Calcule a probabilidade de você ser atendido entre 4 a 5 minutos;
c) Calcule o tempo mínimo em que metade dos clientes serão atendidos;

Exemplo

Considere X como sendo uma variável aleatória que expressa o tempo de espera (em minutos) para que um cliente seja atendido em uma padaria. É conhecido que X segue uma distribuição exponencial com a média de tempo igual a 4 minutos.

a) Esquematize a função densidade de probabilidade de X;
b) Calcule a probabilidade de você ser atendido entre 4 a 5 minutos;

Resposta: 0.0814

c) Calcule o tempo mínimo em que metade dos clientes serão atendidos;

   Resposta: 2,8 minutos

Distribuição normal (gaussiana)

Uma das distribuições mais importantes na estatística (**Teorema Central do Limite**).

Formato de sino;

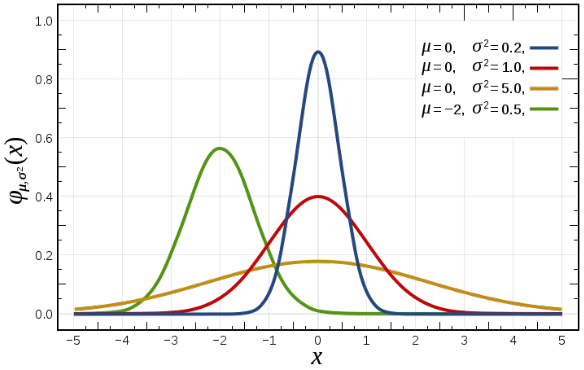
$X \sim N(\mu, \sigma^2)$

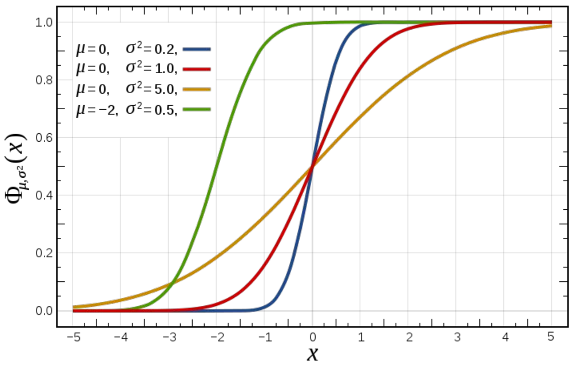
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$F(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx.$$

$$E(X) = \mu$$

$$V(X) = \sigma^2$$





Distribuição normal (gaussiana)

A forma mais simples de uma distribuição normal é quando: $\mu = 0, \sigma^2 = 1$

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

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{x^2}{2}} \, dx.$$

Distribuição normal padrão

Distribuição normal (gaussiana)

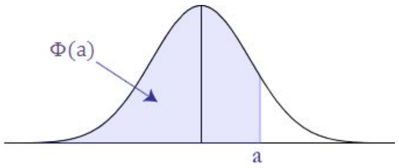
Transformação linear da distribuição normal:

$$X \sim N(\mu, \sigma^2)$$
$$Y = aX + b$$

Y também terá uma distribuição normal!

$$\mu_Y = a\mu_X + b$$
$$\sigma_Y = a\sigma_X$$

$$Z = \frac{X-\mu}{\sigma}$$



Valores de probabilidades para dist. normal padrão ($\mu = 0$, $\sigma^2 = 1$)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Exercício

É conhecido que o nível de glicose no sangue de pacientes diabéticos segue uma distribuição normal de média 106 mg/100 ml e desvio padrão de 8 mg/100 ml.

a) Calcule a probabilidade de uma pessoa diabética aleatória ter um nível de glicose abaixo de 120mg/100 ml.
b) Qual a porcentagem de pessoas diabéticas que possuem o nível de glicose entre 90 e 120 mg/100 ml?

Exercício

É conhecido que o nível de glicose no sangue de pacientes diabéticos segue uma distribuição normal de média 106 mg/100 ml e desvio padrão de 8 mg/100 ml.

a) Calcule a probabilidade de uma pessoa diabética aleatória ter um nível de glicose abaixo de 120mg/100 ml. $P(X \leq 120) = 0.9599$
b) Qual a porcentagem de pessoas diabéticas que possuem o nível de glicose entre 90 e 120 mg/100 ml? $P(90 \leq X \leq 120) = 0.9372 \Rightarrow 93.72\%$

Exercício

É conhecido que o nível de colesterol em homens de 30 anos segue uma distribuição normal com a média 220 mg/dl e desvio padrão de 30 mg/dl. Se existem 20.000 homens com 30 anos na população:

a) Quantos dels possuem colesterol entre 210 e 240 mg/dl?
b) Se um nível de colesterol acima de 250 mg/dl pode provocar trombose, quantos deles possuem o risco de ter trombose?
c) Calcule o nível de colesterol onde 20% dos homens estejam acima dele.

Exercício

É conhecido que o nível de colesterol em homens de 30 anos segue uma distribuição normal com a média 220 mg/dl e desvio padrão de 30 mg/dl. Se existem 20.000 homens com 30 anos na população:

a) Quantos dels possuem colesterol entre 210 e 240 mg/dl? $P(210 \leq X \leq 240) = 0.3781 \Rightarrow 7561.3$
b) Se um nível de colesterol acima de 250 mg/dl pode provocar trombose, quantos deles possuem o risco de ter trombose? $P(X > 250) = 0.1587 \Rightarrow 3173.1$
c) Calcule o nível de colesterol onde 20% dos homens estejam acima dele.

P80=245.2486 mg/dl