

Estatística e probabilidade

🕒 Created	@December 8, 2022 12:25 PM
🏷️ Tags	Study

Conceitos

Variáveis

- Variáveis discretas
São valores distintos entre si, assumem somente valores inteiros
- Variáveis contínuas
São valores contínuos, situados em intervalos $[-\infty, +\infty]$

1. Valor esperado, esperança matemática

$$E(x) = \sum_{i=1}^n X_i \cdot p(X_i)$$

A esperança descreve o valor esperado com base nas probabilidades já conhecidas de um evento acontecer. Então:

Resultados de x	0	1	2	4
Probabilidade	1/8	1/8	1/4	1/2

Nesse caso, a esperança poderia ser calculada por:

$$E(x) = \sum_{i=1}^n X_i \cdot p(X_i) = 0 * \frac{1}{8} + 1 * \frac{1}{8} + 2 * \frac{1}{4} + 4 * \frac{1}{2}$$

Com o resultado sendo igual a

$$E(x) = \frac{21}{8} = 2,625$$

E(x) também pode ser chamado de valor médio das probabilidades = u

Seguindo essa lógica podemos calcular $Var(x)$ da mesma forma que a tradicional

$$Var(x) = \sigma^2 = (x_i - u)^2$$

1.3 Aplicações estatísticas

1.3.1 Cálculo da variância

$$Ex[(x - u)^2]$$

Então seria: $E(x) = \sum_{i=1}^n (xi - u)^2 \cdot p(xi)$

2. Funções de distribuição

2.1 Distribuição acumulada

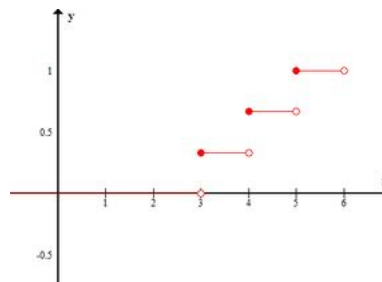
A função de distribuição acumulada se refere a um gráfico de dois eixos, x e y, onde os valores referentes a y descrevem a probabili

A probabilidade de um valor ser menor que zero é sempre igual a zero

$$Fx(X) = 0; x < 0$$

Xi	p(Xi)
X1	0,288
X2	0,432
X3	0,21
X4	0,064

Não devemos pensar nesse gráfico como um que descreve nossa função esperança, ele diz respeito a probabilidade acumulada do



2.1.2 Teoremas:

$$1. Fx(-\infty) = 0; Fx(+\infty) = 1$$

2. A função é sempre não continua a esquerda.

2.1.3 Exemplo

Xi	p(Xi)
0	0.1
4	0.1
5	0.4
8	0.2
10	0.2

▼ Qual a probabilidade de x ser no máximo 5?

$$x \leq 5 \rightarrow p(0) + p(4) + p(5) = 0.1 + 0.1 + 0.4 = 0.6$$

A probabilidade dessa ocorrência é 60%

▼ Qual a probabilidade de x ser no mínimo 7?

Novamente se realiza a soma das probabilidades em que x assuma um valor maior ou igual a 7

Sendo eles $0.2+0.2=0.4$ ou 40%

▼ Qual a probabilidade de x ser menor ou igual a 8 quando x é maior que 4

Esse é um caso de probabilidade condicionada, onde usaremos a equação já conhecida

$$\frac{P(5) + P(8)}{P(5) + P(8) + P(10)}$$

• Respondendo as mesmas questões utilizando a equação a distribuição acumulada

a) No gráfico nós utilizaríamos o valor que a função corresponde em 5

b) Analisariamos o gráfico, iremos pegar o valor em 6(0.6) e em 10(1) e vamos subtrair esse valor.

As perguntas a serem realizadas correspondem a esse formato respondido.

2.2 Distribuição de Poisson

2.2.1 Equação geral

$$P(x) = \frac{\lambda^x \cdot e^{-\lambda}}{x!}$$

Sendo:

λ = O número médio de sucesso num intervalo

x = Número de sucessos

e = número de euler

Podemos usar essa fórmula em valores esperados, ou seja, vamos supor que uma pessoa venda em média 5 produtos por dia, e q

$\lambda = 5$ e $x = 8$. Quando aplicado a probabilidade será igual a 6.58%

▼

O corpo de bombeiros de uma determinada cidade recebe, em média, 3 chamadas por dia. Determine a probabilidade de receber:

a) 4 chamadas por dia

$$\lambda = 3; x = 4$$

$$P(x) = \frac{3^4 \cdot e^{-3}}{4!}$$

Resposta: 16.8%

b) Nenhuma chamada por dia

$$\lambda = 3; x = 0$$

$$P(x) = \frac{\lambda^0 \cdot e^{-\lambda}}{0!} \quad (=) \quad \text{Resposta: 4.9789\%}$$

c) 20 chamadas em uma semana

Note que dessa vez estamos falando de chamadas/semana e não de chamadas/dia, portanto, devemos reajustar nosso λ para semanas, multiplicando por 7 e sendo equivalente a 21. Então iremos prosseguir:

$$P(x) = \frac{21^{20} \cdot e^{-21}}{20!}$$

Resposta: 8.67%

Note que de certa forma essa distribuição busca prever um resultado com base na média já conhecida, ela se preocupa em respon

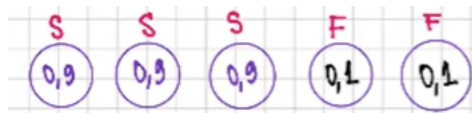
2.3 Distribuição binomial

Imagine que você irá fazer uma pesquisa e os dados que você deseja saber são: o gênero e se essa pessoa gosta de uma comida

Equação da distribuição binomial:

$$P(X = x) = \binom{n}{x} p^x \cdot (1 - p)^{n-x}$$

Como aplicamos essa fórmula? Digamos que você vai realizar 5 provas, mas só conseguiu estudar para 3 delas, e por isso você esi



Iriamos ter: $P(X = 3) = \binom{5}{3} 0.9^3 \cdot (1 - 0.9)^{5-3} = P(X = 3) = \binom{5}{3} 0.9^3 \cdot 0.1^2$

Para relembrar a combinação $\frac{5!}{3!(5-3)!}$

$P(x=3) = 7,29\%$

O valor esperado nessas situações obedece a mesma fórmula de $E(x) = np$ e a variância $Var(x) = np(1 - p)$

Definindo as incógnitas da fórmula: n = número total de tentativas, x= número de sucesso.

▼ Exemplo 1

Uma empresa produz chips de processamento. Em uma de suas grandes fábricas, 2% dos chips produzidos estão com algum tip

a) Média

A média em uma distribuição binomial é a mesma coisa que o valor esperado $E(x) = np$

$$E(x) = 500 \cdot 0,02 = 10$$

b) Desvio padrão

$$\sigma = \sqrt{Var(x)} = \sqrt{np(1 - p)} = \sqrt{500 \cdot 0,02(1 - 0,02)} = 3,13$$

2.4 Distribuição normal

A distribuição normal é a distribuição mais comum de ser vista em apresentações ou relatórios de negócios, isso se deve, porque a

- O ponto máximo da função corresponde à média.
- O máximo é o valor mais provável de ser encontrado no conjunto.
- Os valores ao lado do máximo têm probabilidade sempre inferior.
- A probabilidade corresponde a área.
- A probabilidade está distribuída de forma 50/50 em torno da média

Normalização:

Também chamado de escore-z (pode ser usado para os outros casos de distribuição)

$$z = \frac{x - u}{\sigma}$$

Devemos utilizar a normalização para não ter que fazer integrais em toda tentativa.

→ Perguntas do tipo “Qual a chance de um evento ocorrer se x tentativas forem feitas em n vezes?” Na fórmula z, o número n ocupa

Exemplo:

— Se extrairmos de um lote de 100 de peças ao acaso. Sabendo que a probabilidade de se obter peças defeituosas é de 10%. (

$$\mu = n.p = 100.0,1 = 10$$

$$\text{Var}(x) = np(1 - p) = 100.0,1.0,9 = 9$$

$$\sigma = \sqrt{9} = 3$$

Nesse caso, a distribuição normal tem seu ponto máximo quando $x=10$ e desejamos saber quando menos que 20 são defeitu

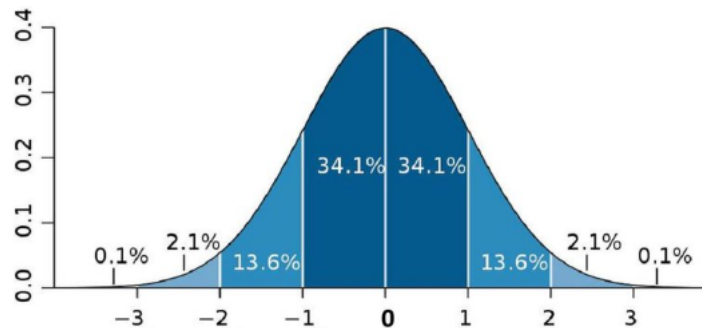
$$\int_{-\infty}^{20} f(x)$$

, ou como sabemos que o que está a esquerda do ponto máximo equivale a 50%, então teremos que calcular apenas o esco

Olharemos o valor na nossa tabela-z para 3,33 e pronto. Faremos a soma desse valor e 50%.

Equação:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2} \cdot \left(\frac{x-u}{\sigma}\right)^2}$$



Exemplos:

2.4.1 Suponha que entre investidores, o ganho entre uma aplicação financeira tenha uma distribuição aproximadamente normal (

Organizando: $u = 105.000,00$; $\sigma = 10.000,00$; $P(90.000,00 < x < 125.000,00)$

Usando a normalização:

$$P(z_1 < x < z_2) = \left(\frac{90 - 105}{10} < x < \frac{125 - 105}{10} \right)$$

$$P(-1,50 < x < 2,00)$$

Mas agora, o que esses números significam? Esses números são tabelados e devem ser consultados levando em consideração z_n

Aproximação da binomial pela normal

3. Função de densidade de probabilidade

A função é caracterizada por:

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

Define a probabilidade de ocorrência de um evento ao longo de um gráfico. É muito simples entender esse conceito, usando o conh

- $f(x) \geq 0$
- $P(a \leq x \leq b) = \int_a^b f(x)dx$

4. Probabilidade conjunta de variável aleatória bidimensional

Nós sabemos que o mundo real é bem mais complexo que uma análise dimensional simples, e toda decisão ou conclusão depende

→ x = preço

→ y = quantidade de comida

Vale lembrar que você está com fome e possui apenas 25 reais a serem gastos. Dessa maneira podemos organizar uma tabela

$x = [21, 23, 24]$

$y = [0.320, 0.280, 0.400, 0.420]$ (g)

x/y	0.320	0.280	0.400	0.420	P(X=xi)
21	0,04	0,05	0,01	0	0,1
23	0,08	0,15	0,04	0,03	0,30
24	0,06	0,14	0,20	0,20	0,6
P(Y=yi)	0,18	0,34	0,25	0,23	1

A partir disso nós podemos ter uma noção melhor de como funciona a probabilidade de ocorrência bidimensional