Introdução à Incerteza de medição

NOÇÕES DE PROBABILIDADE

AULA 02

REALIZAÇÃO

sociedade brasileira de metrologia

Sumário

Apresentação

Olá! Seja muito bem-vindo à segunda aula sobre Introdução à Incerteza da Medição.

Na aula de hoje falaremos sobre um assunto que usamos com frequência em nosso cotidiano: probabilidade e distribuição de probabilidade.

Ao final dessa aula, serão disponibilizados exercícios para fixação, lembre-se de fazê-los, pois assim você poderá verificar se realmente compreendeu o assunto trabalhado nessa aula.

Bons estudos!

1. Introdução à probabilidade

A Metrologia não é uma ciência absoluta, e sim uma ciência probabilística. Ela se baseia em probabilidades de ocorrência e se utiliza da estatística como ferramenta de análise. Daí a importância do conhecimento de conceitos estatísticos básicos.

Você já apostou alguma vez na Mega-Sena? Sabe qual é sua chance de ganhar? Para quem nunca apostou, nesse concurso os apostadores escolhem 6 números, em uma cartela com dezenas de 1 a 60. São sorteadas 6 dezenas, e quem acertar as seis dezenas ganha, normalmente, alguns milhões de reais.

Segundo a Caixa Econômica Federal (CEF), a chance de ganhar, apostando seis dezenas, é de uma em 50 milhões. Sim, a chance é bem pequena. Se você quiser, a CEF permite jogar até 15 dezenas na mesma cartela, e sua chance de ganhar passa a ser uma em 10 mil, ou seja, você terá 5000 vezes mais chance de acerto (mas vou avisando, sua aposta ficará 5000 vezes mais cara!).

Certo, mas, como chegamos nesses valores? Como é que descobrimos a chance de um determinado evento acontecer?

É aqui que entra o conceito de probabilidade. Se quisermos saber a chance de algo acontecer devemos, primeiro, descobrir quantas são as possibilidades para aquele evento ocorrer.

Quer um exemplo? Vamos a ele:

Maria ficou grávida de gêmeos, mas ela ainda não sabe o sexo dos bebês. Supondo que a chance de cada bebê ser menino ou menina é a mesma, vamos ver quais são as possibilidades de sexo para cada um eles?

Para facilitar o entendimento vamos pensar na ordem de nascimento deles, ok?
Então vejamos:

POSSIBILIDADE	1° BEBÊ A NASCER	2° BEBÊ A NASCER
1	Menino	Menino
2	Menina	Menino
3	Menino	Menina
4	Menina	Menina

Temos 4 possibilidades, duas onde Maria terá um casal, uma onde ela terá dois meninos e uma onde ela terá duas meninas.

Portanto, se quisermos saber a chance de ela ter dois meninos, devemos dividir o número de possibilidades onde se tem dois meninos pelo número de possibilidades total. Logo:

$$Chance\ de\ ter\ 2\ meninos = \frac{1}{4} = 0{,}25 = 25\%$$

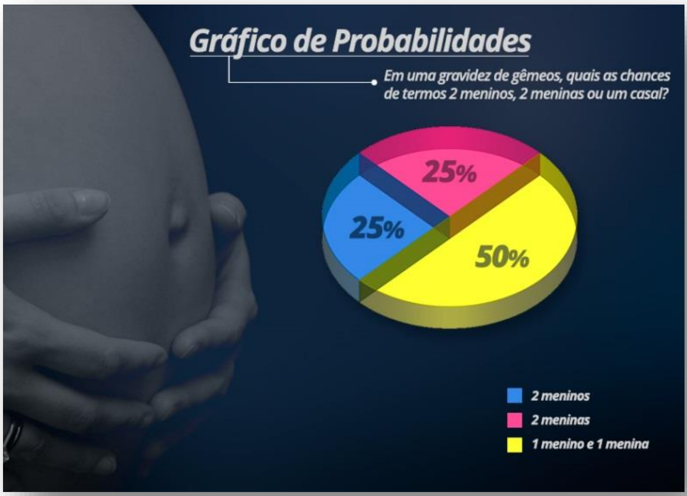
De forma análoga, as chances de termos duas meninas ficam assim:

$$Chance\ de\ ter\ 2\ meninas = \frac{1}{4} = 0{,}25 = 25\%$$

Agora, se Maria quiser ter um casal, as chances são um pouco maiores, já que temos duas possibilidades onde isso ocorre. Logo, a chance dela ter um casal será dada por:

$$Chance\ de\ ter\ 1\ menino\ e\ 1\ menina = \frac{2}{4} = 0{,}50 = 50\%$$

Observe:



O mesmo ocorre para qualquer outro evento.

Vamos voltar ao caso da Mega-Sena.

Quantas são as combinações possíveis de 6 dezenas em 60 números a escolher (de 1 a 60)?

O cálculo é um pouco complicado, mas o número final será esse:

$$50.063.860$$

Mais de 50 milhões de possibilidades. Se você fizer apenas um jogo, a sua chance de ganhar é:

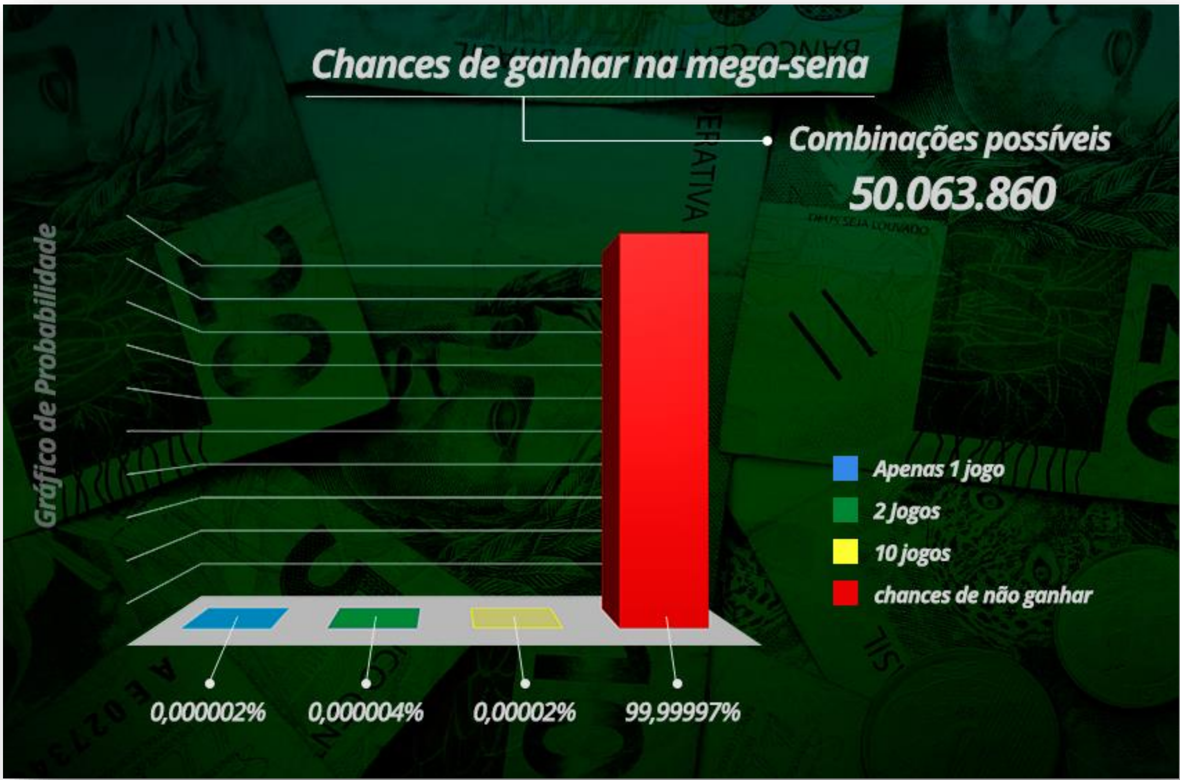
$$Chance\ de\ ganhar\ na\ Mega\ Sena = \frac{1}{50.063.860} \approx 0{,}00000002 = 0{,}000002\%$$

Agora, se você apostar em 2 combinações diferentes:

$$Chance\ de\ ganhar\ na\ Mega\ Sena = \frac{2}{50.063.860} \approx 0{,}00000004 = 0{,}000004\%$$

E se você apostar 10 combinações:

$$Chance\ de\ ganhar\ na\ Mega\ Sena = \frac{10}{50.063.860}$$
$$\approx 0{,}0000002 = 0{,}00002\%$$



Se a sua chance de ganhar tivesse o comprimento de uma fita de 80 cm, sua chance de não ganhar tem o comprimento de uma fita que poderia dar a volta na Terra.

Se imaginarmos que a chance de não ganhar na Mega-Sena é do tamanho do Everest, a chance de alguém ganhar é mais ou menos o tamanho de um palito de fósforo, deitado.

Então a chance é bem pequena mesmo...

Segundo o site Mega Curioso, é mais fácil você ser mordido por uma cobra venenosa (1 chance em 3,5 milhões) ou ser atingido por um raio (1 em 1 milhão), do que ganhar na Mega-Sena. Na verdade, é bem mais fácil você ganhar uma medalha olímpica (1 chance em 662 mil) do que ganhar o prêmio da Mega-Sena.

Mas fique tranquilo, a chance da sua casa ser destruída por um meteoro é bem menor, é de apenas 1 em 182 trilhões. Ufa!
Bom, mas continuando...

Função distribuição de probabilidades

Entende-se por função distribuição de probabilidade a relação entre uma determinada variável aleatória **x** e a probabilidade de sua ocorrência **p(x)**.

Existem diversas funções de distribuição de probabilidades, dentre as quais podemos destacar as mais utilizadas na metrologia que são a distribuição uniforme, a triangular, a normal e a t-Student.

Independente da denominação, a função distribuição de probabilidades associa uma probabilidade a cada resultado numérico de um experimento, fornecendo a probabilidade de ocorrência em maior ou menor extensão para cada valor da variável aleatória, seja ela discreta ou contínua.

Primeiro vamos definir o que vem a ser uma **função**.

Uma **função matemática** nada mais é do que uma lei que estabelece **uma relação entre uma variável independente**, vamos chamá-la de ***x***, **e uma variável dependente**, que neste caso vamos chamar de ***y***. Bom, se y depende de x, toda vez que alterarmos o valor de **x**, o valor de **y** também mudará, afinal, y é uma variável dependente de x.

Quer ver um exemplo?

Wesley é garçom em um restaurante, e ele sempre recebe, além de seu salário, 10% sobre o valor da conta de cada cliente que ele atende durante o dia. É o que chamamos de gorjeta, ou taxa de serviço. O valor da gorjeta sempre **depende** do valor da conta do cliente, portanto podemos dizer que o valor da gorjeta é uma função do valor da conta do cliente.

Se colocarmos em fórmula matemática:

$$valor\ da\ gorjeta = \frac{conta\ do\ cliente}{10}$$

Se chamarmos o valor da conta do cliente de **x** (variável independente) e o valor da gorjeta de **y** (variável dependente), então teremos:

$$y = \frac{x}{10}$$

Para deixar mais explícita a relação entre **x** e **y**, usamos a notação **f(x)** para designar a variável dependente **(y)**.

Então fica assim:

$$f(x) = \frac{x}{10}$$

Um recurso prático que podemos usar com funções matemáticas é a construção de gráficos.

Com um gráfico, podemos analisar o comportamento de uma função de forma mais clara, ver como é sua evolução em função de sua variável independente. Se plotarmos a função que define o valor da gorjeta, teremos o seguinte gráfico:

Plotar
Termo utilizado para o desenho de gráficos.

Para "ler" o gráfico é muito simples. Ache o valor da conta no eixo horizontal (eixo **x**) e trace uma linha vertical até interceptar a curva do gráfico. Marque esse ponto e então trace uma linha horizontal a partir desse ponto até "bater" no eixo vertical (eixo **y**). Dessa forma você sabe automaticamente o valor da função

y quando escolhido um valor para a variável independente **x**.

Abaixo mostramos um exemplo no qual o cliente gastou R$ 2000,00 reais na conta do restaurante.

Observe:

Perceba que com a ajuda do gráfico, pudemos descobrir o valor da gorjeta de Wesley, que nesse caso, foi de R$ 200,00.

Voltando a questão das probabilidades...

A probabilidade da ocorrência de um determinado evento também pode ser descrita na forma de uma função.

Lembra do exemplo da Mega-Sena?

Podemos dizer que a probabilidade de você ganhar na Mega-Sena é uma função que depende da quantidade de combinações que você apostar. Quanto maior o número de combinações jogadas (**x**) maior é sua chance de ganhar (**f(x)**).

Vamos ver como fica no gráfico?

O que temos acima é a distribuição de probabilidade de uma pessoa qualquer (como eu e você) de acertar na Mega-Sena.

Reforçando: a função que descreve a probabilidade de um evento qualquer ocorrer é denominada função distribuição de probabilidade.
Esse conceito é de fundamental importância para muitas áreas do conhecimento, como a física, geografia, ciências da computação e também para a metrologia.

Mas não se apresse, ainda precisamos aprofundar um pouco mais nossos conhecimentos sobre distribuições de probabilidade.

Acompanhe!

2. Distribuição discreta

A Distribuição discreta descreve quantidades aleatórias que podem assumir valores particulares finitos.

Vamos pensar na probabilidade de ganharmos no cara ou coroa como uma função de probabilidades, nesse caso, a variável **x** aleatória é a face da moeda, e a probabilidade de sua ocorrência que chamaremos de **p(x)** é a função de distribuição de probabilidades.

Vejamos:

Se eu escolher cara, e você escolher coroa, qual a chance que você tem de ganhar ao jogarmos a moeda para cima?
E qual é a minha chance de ganhar?
Se a moeda for honesta, nós dois teremos a mesma chance de ganhar, logo, 50% de chances para cada um.

Você vence: 50%
Eu venço: 50%

Certo, agora pense comigo, será que existe alguma chance de, jogarmos a moeda e nenhum de nós vencer?

Isso mesmo! Não há essa possibilidade, assim como também não há a chance de vencermos no jogo simultaneamente.

Logo:
Nenhum vence: 0%
Os dois vencem: 0%

Portanto, se quisermos descrever o cara ou coroa como uma função de distribuição de probabilidades teremos o seguinte:

$$p(x) = 0{,}5 = \frac{1}{2} = 50\%$$

Onde **x** pode assumir apenas dois valores possíveis: cara, ou coroa.

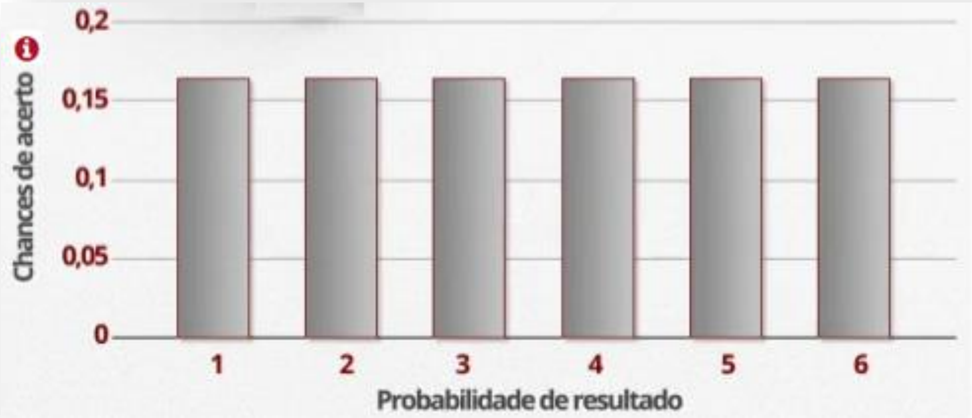
Toda vez que uma determinada variável **x** puder assumir apenas certos valores pré-determinados, dizemos que a distribuição de probabilidades de **x**, a **p(x)**, é uma **distribuição de probabilidade discreta**.

Vamos a um segundo exemplo:

Imagine um dado e os valores que ele pode assumir, considere esses valores, que vão de 1 a 6, como uma variável aleatória **x**. Teremos, então que: **x** = (1; 2; 3; 4; 5; 6). Ao lançarmos este dado, a probabilidade de x assumir qualquer um dos valores é de 1 para 6, (um sexto), ou seja:

$$p(x) = \frac{1}{6}$$

Observe que ao lançarmos um dado não podemos encontrar, por exemplo, valores como 1,4 ou 5,75. Assim, a distribuição de probabilidade **p(x)** desta **variável é discreta** porque a variável **não pode assumir valores intermediários**.



Observe o gráfico a seguir:

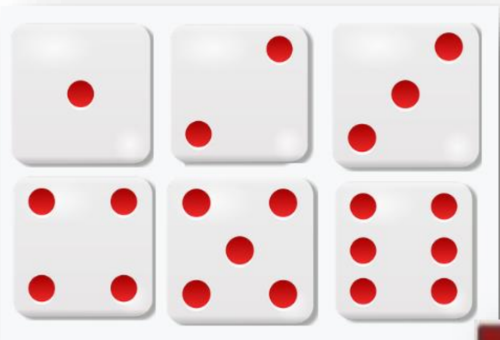
> *Apenas para esclarecer...*
>
> *Nas chances de acerto colocamos as probabilidades em valores decimais...*
> *Onde você lê 0,2 é o mesmo que ler 20%, onde você lê 0,15 é o mesmo que ler 15% e assim por diante.*

Note também que todos os valores têm a mesma chance de ocorrer, ou seja, não há um valor favorito ou mais provável. Todos têm 16,67% de chance.

Agora vamos fazer um exercício de imaginação e alterar o nosso dado.

Primeiro imagine um dado normal...

Ele é assim:



Valor	Núm. de faces
●	1
●●	1
●●●	1
●●●●	1
●●●●●	1
●●●●●●	1

Agora vamos fazer o seguinte:

Valor	Núm. de faces
● ● ●	1
● ●	1
● ● ●	1
● ● ● X	1
● ● ● ● ●	1
● ● ● ● ● ●	1

➡ Pintar mais duas bolinhas na face que só tem uma bolinha

➡ Retirar uma bolinha da face que tem 4 bolinhas

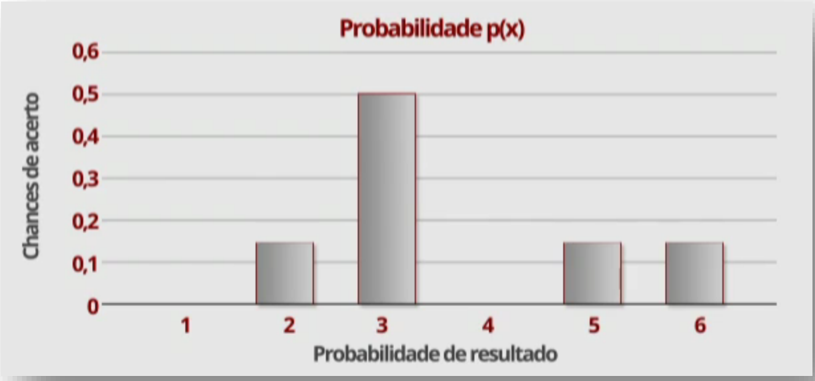
Então nosso dado alterado ficará assim:

Valor	Núm. de faces
●	0
● ●	1
● ● ●	3
● ● ● ●	0
● ● ● ● ●	1
● ● ● ● ● ●	1

Agora vamos escrever as chances de cada valor ocorrer nesse caso?

Valor	Núm. de faces	Probabilidade
●	0	0/6
● ●	1	1/6
● ● ●	3	3/6
● ● ● ●	0	0/6
● ● ● ● ●	1	1/6
● ● ● ● ● ●	1	1/6

Vamos montar o gráfico novamente:



Você percebeu que agora temos um "favorito"?

Exato! É o número 3...

Se fôssemos apostar, certamente apostaríamos no número 3, porque ele tem 50% de chances de cair. Depois os números 2, 5 e 6, que tem 16,67% de chances. Os números 1 e 4 são piores ainda, pois tem 0% de chance de sair, portanto, se quiser alguma chance de ganhar, não aposte neles...

Vamos ao próximo conceito?

3. Distribuição contínua

A distribuição contínua descreve as probabilidades dos possíveis valores de uma variável aleatória contínua. Uma variável aleatória contínua possui um conjunto de valores infinito e incalculável.

As probabilidades de variáveis aleatórias contínuas (**x**) são definidas como a área sob a curva da sua distribuição. Assim, apenas as faixas de valores podem ter uma probabilidade diferente de zero. A probabilidade de que uma variável aleatória contínua seja igual a algum valor é sempre zero.

E o que isso significa?

Nem sempre uma variável irá assumir apenas um número limitado de possibilidades. Pense no seu dia a dia, na conta de luz, por exemplo. A energia elétrica está sendo medida a todo o instante. Toda casa, dependendo da época do ano, tem um consumo médio de energia elétrica, que pode variar um pouco de um mês para o outro. Se gastarmos mais energia que o habitual, tomando banhos mais demorados ou usando mais o ferro de passar, a conta virá acima da média, mas se tomarmos medidas simples como apagar a luz toda vez que sair de um cômodo, ou desligar aparelhos que não estão sendo utilizados, o valor da conta será abaixo da média.

Mas quanto será a variação da conta, em reais? Não sei, depende do quanto você economizar ou não. Pode ser alguns centavos ou até dezenas de reais, dependendo das medidas que você tomar.

Imagine o seguinte:

Se uma pessoa tem, em sua casa, um consumo médio de energia elétrica de 200 kWh por mês, sua conta de luz dará mais ou menos uns R$ 138,00.

Se pensarmos apenas no valor da conta, podemos dizer que 138 reais é a média do nosso exemplo.

Portanto, se fôssemos tentar adivinhar o valor da próxima conta de luz dessa pessoa, certamente R$ 138,00 seria o melhor "chute" que poderíamos dar.

Certo, mas agora pense, em condições normais, o valor da conta dessa pessoa pode vir, por exemplo, R$ 138,42 em um mês e R$ 137,58 no outro. A média ainda será 138 nos dois meses, e esse valor, mesmo que não tenha aparecido "exato", ainda será o valor mais provável para a conta dela.

Nesse caso, o valor 138 é a média, ou seja, o valor isolado que tem maior chance de aparecer, mesmo ele nunca tenha aparecido com exatidão.

Agora pense no valor R$ 85,00, você acha que ele tem grande ou pequena chance de aparecer na conta de luz dessa mesma pessoa? E o R$ 956,00? E o R$ 4,75? E o R$ 1458,00?

A chance de vir um valor desses é muito pequena porque está muito longe da realidade cotidiana dessa pessoa. A menos que ela compre um secador de cabelo industrial, ou passe a iluminar a casa apenas com velas, ou que a empresa de energia elétrica aumente a tarifa em 10 vezes, não existe muitas chances de ela ter uma conta de luz muito diferente de 138 reais.

Agora vamos ver como ficaria a distribuição de probabilidades, supondo um **desvio padrão** de 20 reais.

> ***Lembrando:*** *O* ***desvio padrão*** *é a medida de dispersão mais importante para a Metrologia. Com essa medida, podemos ter uma* ***noção precisa*** *da variação dos valores em torno da média. Basicamente,* ***quanto menor for o desvio padrão****, menor será a dispersão dos valores, ou seja,* ***maior será o grau de precisão dessa medida.***

média. Sempre que nos afastamos do valor central, para mais ou para menos, o valor da probabilidade, **p(x)**, diminui drasticamente.

Lembre-se também que qualquer valor sempre será possível, logo, valores muito distantes da média tem uma probabilidade pequena, portanto maior que zero, de ocorrer. É o mesmo que se perguntar "há alguma chance de eu achar um pote de ouro perdido?" Há sim! A chance é bem pequena, MAS ELA EXISTE. Portanto siga tentando...

E se quisermos representar matematicamente isso ficaria assim: **p(0)>0** assim como a **p(+∞)>0**

Vamos a mais um exemplo.

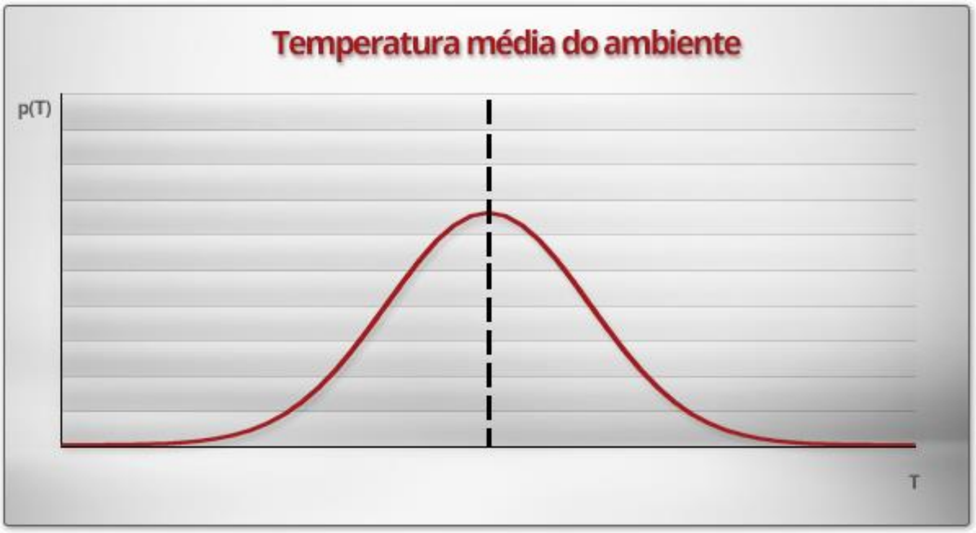
Imagine uma situação, na qual a temperatura ambiente de um laboratório vem sendo monitorada e medida ao longo de um dia.

Normalmente, a temperatura é mais baixa pela manhã e vai aumentando progressivamente até estabilizar.
Conforme a tarde avança e o sol vai baixando, a temperatura vai diminuindo até o anoitecer.

Digamos que pela manhã a temperatura estivesse em 18°C e durante a tarde ela chegasse a 22°C, isso dá uma média de 20°C durante esse dia. Então poderíamos ter, ao longo do dia, qualquer temperatura entre 18°C e 22°C, como por exemplo 19,2°C ou 21,5°C, mas a temperatura com maior probabilidade continuaria sendo 20°C, pois esse é o valor central, ou seja, a média encontrada entre as temperaturas máximas e mínimas registradas durante o dia.

Visto isso, você concorda que a temperatura ambiente é uma variável contínua, pois ela pode assumir qualquer valor ao longo do dia? Desta forma, a distribuição de probabilidade também será contínua.

Se traçarmos uma curva de distribuição dos valores medidos e considerarmos que a temperatura média do laboratório é o valor central desta distribuição, a função distribuição de probabilidade da temperatura terá a seguinte forma:



Nesse caso, a temperatura média do ambiente é o valor central dessa distribuição, e, portanto, é o valor com maior probabilidade de ocorrência. À medida que nos afastamos da temperatura média, para mais ou para menos, a probabilidade **p(T)** diminui de forma contínua até valores muito pequenos.

Observação

Devemos notar uma importante propriedade de qualquer função distribuição de probabilidade **p(x),** seja ela discreta ou contínua:

$P(-\infty < x < +\infty) = 1$

Probabilidade da função P de x
A probabilidade da função P de x, no intervalo que vai de menos infinito a mais infinito, é igual a um.

Isto é, o somatório dos valores de **p(x)**, no caso de distribuição discreta, ou a integral de **p(x)**, no caso de distribuição contínua, entre os limites - ∞ e + ∞, que corresponde à probabilidade de **x** estar dentro dos limites, sempre resulta em 1.

Na prática, isso significa dizer que, a chance de ocorrer qualquer valor de **x**, entre menos e mais infinito será de 100%. É como se você pudesse apostar em todos os números de uma roleta, ou em todos os cavalos em uma corrida, sua chance de ganhar será de 100%.

No exemplo do dado não viciado, o somatório de p(x) será:

$$p(-\infty < x < +\infty) = p(1 \leq x \leq 6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = 1$$

Soma das probabilidades
A soma das probabilidades de ocorrência de cada uma das faces (de 1 a 6) é igual a um.

No caso da temperatura do laboratório, a **área abaixo da curva**, que representa a integral da função, também valerá 1, ou seja:

$$\int_{-\infty}^{+\infty} p(T)dT = 1$$

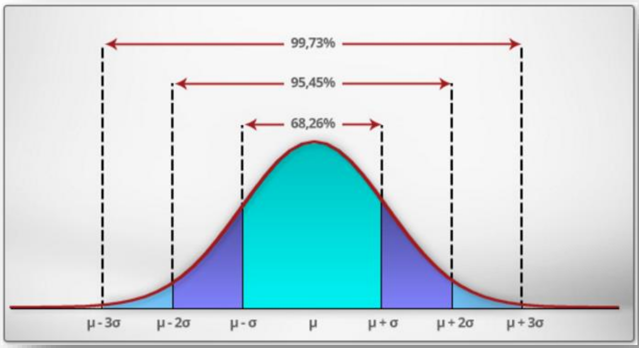
Integral da função P de T
A integral da função p de T, no intervalo de menos infinito a mais infinito, em relação a T é igual a um.

Entendido?

Como já informamos, as distribuições contínuas de probabilidades mais utilizadas na Metrologia são a Normal ou Gaussiana, Uniforme ou Retangular, Triangular e t-Student.

Vamos a elas:

4. Distribuição normal



A distribuição Normal ou Gaussiana é, sem dúvida, a mais importante das distribuições de probabilidade dentre as usadas na Metrologia, pois além de descrever uma série de fenômenos físicos e financeiros, possui grande uso na estatística inferencial.

A distribuição Normal é inteiramente descrita por seus parâmetros de média e desvio padrão, isso significa que, conhecendo estes valores consegue-se determinar qualquer probabilidade em uma distribuição Normal.

Como você pode observar na imagem ao lado, a função densidade de probabilidade p(x) da distribuição normal tem a forma de um sino.

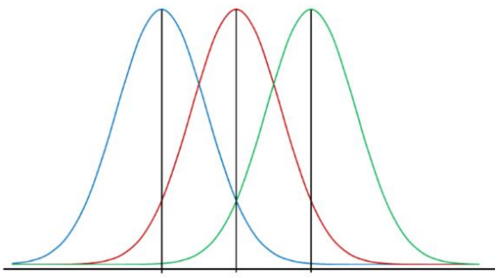
***Curiosidade**:*

Johann Carl Friedrich Gauss (1777 – 1855): Matemático, astrônomo e físico, contribuiu em diversos campos da ciência, dentre eles a teoria dos números, estatística, análise matemática e geometria diferencial. Desenvolveu o método dos mínimos quadrados aos dezoito anos de idade, trabalho que despertou o interesse de Gauss na teoria dos erros de observação, que culminou na lei de Gauss da distribuição normal de erros e sua curva em formato de sino, que hoje é comumente utilizada por todos aqueles que trabalham com estatística.

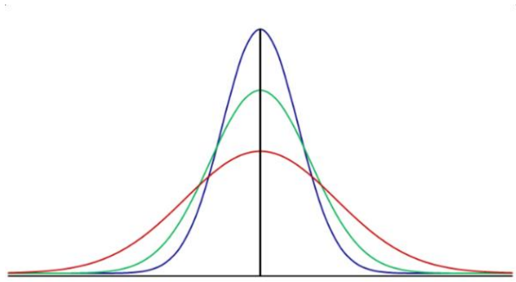
O valor central, μ, representa a **média**, e σ representa o **desvio padrão**. Pode-se se notar que, partindo-se da média, ao nos deslocarmos um desvio padrão, para mais e para menos, temos um intervalo que engloba 68,26% da população de valores possíveis. Se nos deslocarmos dois desvios padrão, o intervalo terá uma abrangência de 95,45%, e se o intervalo for de $\pm 3\sigma$, teremos 99,73% dos valores próximos à média.

Outra curiosidade: na distribuição normal o desvio padrão representa o ponto de inflexão da curva, ou seja, o ponto em que a curva muda sua inclinação.

Como dito, o que caracteriza uma distribuição normal é a média (μ) e o desvio padrão (σ). Diferentes combinações média/desvio padrão correspondem a formatos de curvas diferentes.



Médias diferentes e desvios padrões iguais



Médias iguais e desvios padrão diferentes

Veja a expressão matemática da distribuição normal:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad \infty < \mu < +\infty \text{ e } \sigma > 0$$

Onde

e = 2,718 (constante)

π = 3,1416 (constante)

μ = média da população

σ = desvio padrão da população

x = variável aleatória

Lembrando que essa função descreve a curva normal, logo, se integrarmos essa função teremos a área abaixo da curva, que representa o grau de abrangência do intervalo selecionado.

Quanto maior o número de medições feitas de um mesmo mensurando, mais próximos de uma curva normal os seus valores ficarão. Assim, se forem realizadas infinitas medições chegaremos a uma distribuição normal.

Traduzindo em miúdos: Se você fizer uma pesquisa com duas ou três pessoas, terá uma distribuição, mas se você for aumentando o número de entrevistados, quando chegar na casa dos milhares, terá uma curva normal em suas mãos.

Mas você deve estar se perguntando "Nossa! Preciso mesmo entender essa equação?”

Não se preocupe, pois, o objetivo aqui não é esse.

A utilização direta dessa Equação torna-se matematicamente complexa, mas por intermédio de um modelo matemático simplificado criou-se a **distribuição normal padronizada,** dada pela seguinte equação:

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

A função p de x
A função p de x é igual a 1 sobre sigma vezes a raiz quadrada de dois pi, multiplicado por e elevado a menos z ao quadrado sobre dois.

Onde $z = \frac{x-\mu}{\sigma}$, μ é a média e σ é o desvio padrão.

Na distribuição normal padronizada μ assume valor zero e σ valor um.

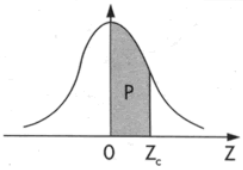
Nos dados originais a média e o desvio padrão podem assumir diferentes valores, entretanto, utilizando-se apenas uma tabela, qualquer variável ou dado é transformado na variável **Z,** cujo valor é a variável menos a média dividido pelo desvio padrão.

Veja abaixo a tabela de distribuição normal:

Tabela III — Distribuição Normal Padrão

$Z \sim N(0, 1)$

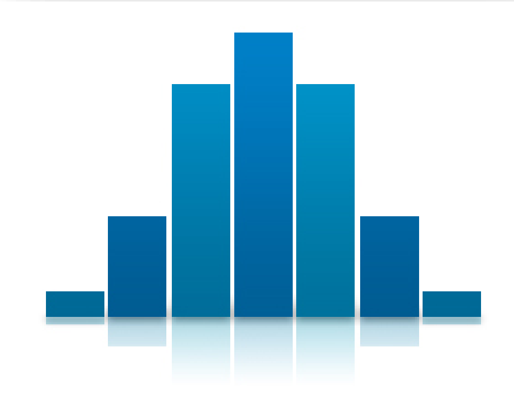
Corpo da tabela dá a probabilidade p, tal que $p = P(0 < Z < Z_c)$



parte inteira e primeira decimal de Z_c	Segunda decimal de Z_c: 0	1	2	3	4	5	6	7	8	9	parte inteira e primeira decimal de Z_c
	p = 0										
0,0	00000	00399	00798	01197	01595	01994	02392	02790	03188	03586	0,0
0,1	03983	04380	04776	05172	05567	05962	06356	06749	07142	07535	0,1
0,2	07926	08317	08706	09095	09483	09871	10257	10642	11026	11409	0,2
0,3	11791	12172	12552	12930	13307	13683	14058	14431	14803	15173	0,3
0,4	15542	15910	16276	16640	17003	17364	17724	18082	18439	18793	0,4
0,5	19146	19497	19847	20194	20540	20884	21226	21566	21904	22240	0,5
0,6	22575	22907	23237	23565	23891	24215	24537	24857	25175	25490	0,6
0,7	25804	26115	26424	26730	27035	27337	27637	27935	28230	28524	0,7
0,8	28814	29103	29389	29673	29955	30234	30511	30785	31057	31327	0,8
0,9	31594	31859	32121	32381	32639	32894	33147	33398	33646	33891	0,9
1,0	34134	34375	34614	34850	35083	35314	35543	35769	35993	36214	1,0
1,1	36433	36650	36864	37076	37286	37493	37698	37900	38100	38298	1,1
1,2	38493	38686	38877	39065	39251	39435	39617	39796	39973	40147	1,2
1,3	40320	40490	40658	40824	40988	41149	41309	41466	41621	41774	1,3
1,4	41924	42073	42220	42364	42507	42647	42786	42922	43056	43189	1,4
1,5	43319	43448	43574	43699	43822	43943	44062	44179	44295	44408	1,5
1,6	44520	44630	44738	44845	44950	45053	45154	45254	45352	45449	1,6
1,7	45543	45637	45728	45818	45907	45994	46080	46164	46246	46327	1,7
1,8	46407	46485	46562	46638	46712	46784	46856	46926	46995	47062	1,8
1,9	47128	47193	47257	47320	47381	47441	47500	47558	47615	47670	1,9
2,0	47725	47778	47831	47882	47932	47982	48030	48077	48124	48169	2,0
2,1	48214	48257	48300	48341	48382	48422	48461	48500	48537	48574	2,1
2,2	48610	48645	48679	48713	48745	48778	48809	48840	48870	48899	2,2
2,3	48928	48956	48983	49010	49036	49061	49086	49111	49134	49158	2,3
2,4	49180	49202	49224	49245	49266	49286	49305	49324	49343	49361	2,4
2,5	49379	49396	49413	49430	49446	49461	49477	49492	49506	49520	2,5
2,6	49534	49547	49560	49573	49585	49598	49609	49621	49632	49643	2,6
2,7	49653	49664	49674	49683	49693	49702	49711	49720	49728	49736	2,7
2,8	49744	49752	49760	49767	49774	49781	49788	49795	49801	49807	2,8
2,9	49813	49819	49825	49831	49836	49841	49846	49851	49856	49861	2,9
3,0	49865	49869	49874	49878	49882	49886	49889	49893	49897	49900	3,0
3,1	49903	49906	49910	49913	49916	49918	49921	49924	49926	49929	3,1
3,2	49931	49934	49936	49938	49940	49942	49944	49946	49948	49950	3,2
3,3	49952	49953	49955	49957	49958	49960	49961	49962	49964	49965	3,3
3,4	49966	49968	49969	49970	49971	49972	49973	49974	49975	49976	3,4
3,5	49977	49978	49978	49979	49980	49981	49981	49982	49983	49983	3,5
3,6	49984	49985	49985	49986	49986	49987	49987	49988	49988	49989	3,6
3,7	49989	49990	49990	49990	49991	49991	49992	49992	49992	49992	3,7
3,8	49993	49993	49993	49994	49994	49994	49994	49995	49995	49995	3,8
3,9	49995	49995	49996	49996	49996	49996	49996	49996	49997	49997	3,9
4,0	49997	49997	49997	49997	49997	49997	49998	49998	49998	49998	4,0
4,5	49999	50000	50000	50000	50000	50000	50000	50000	50000	50000	4,5

Continuando...

5. Teorema do limite central



Existe um teorema, denominado teorema do limite central, que demonstra que:

Qualquer que seja a distribuição da variável de interesse, a distribuição das médias amostrais tenderá para uma distribuição normal à medida que o tamanho de amostra cresce.

Pode-se ter uma variável original com uma distribuição muito diferente da normal, pode até mesmo ser discreta, mas se tomarmos várias amostras desta distribuição e fizermos um **histograma** das médias amostrais, a forma tenderá a uma curva normal. Aumentando-se o tamanho das amostras as médias amostrais tenderão à uma distribuição normal. Assim,

Histograma
Histograma é a representação gráfica, em colunas (retângulos), de um conjunto de dados previamente tabulado e dividido em classes uniformes.

qualquer amostra retirada da população aproxima-se da distribuição normal com média$(\bar{x})$ e desvio padrão s_x.

O desvio padrão das médias das amostras $(s_{\bar{x}})$, chamado de **erro padrão** de estimativa, é calculado da seguinte forma:

$$s_{\bar{x}} = \frac{s_x}{\sqrt{n}}.$$

O teorema do limite central funciona bem a partir de amostras com n ≥ 30. Entretanto, trabalhamos muitas vezes com amostras inferiores a 10. Assim, na metrologia, a estimativa da média da população é feita com base na distribuição t –Student, que é derivada da distribuição normal padronizada.

E próxima distribuição é:

6. Distribuição t-Student

Curiosidade:

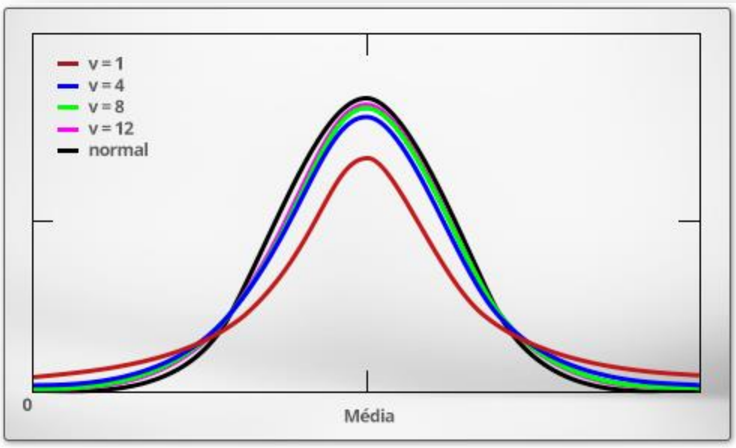
William Sealy Gosset (1876-1937): *Químico e matemático, trabalhou para Arthur Guinness de 1899 até 1935 na cervejaria Guinness. Na época, Guinness buscava cientistas que pudessem apresentar suas habilidades para o processo de fabricação de cerveja, e em 1904 Gosset escreveu um relatório interno intitulado "A aplicação da Lei do Erro", no qual introduziu várias metodologias estatísticas para a indústria cervejeira. O primeiro artigo de Gosset foi uma aplicação da distribuição de Poisson na contagem de leveduras. Para muitos usuários da Estatística o nome William Gosset soa desconhecido, mas o pseudônimo* ***Student*** *o revela como um dos estatísticos mais importantes e brilhantes da história.*

A distribuição t é uma distribuição de probabilidade teórica. É simétrica, em forma de sino, e semelhante à curva normal padrão, porém com caudas mais largas. Uma simulação da t-Student pode gerar valores mais extremos que uma simulação da normal.

Como vimos anteriormente, é necessário um grande número de medições para se obter uma distribuição próxima à normal.

Como nem sempre é viável realizar 30 medições de um mesmo mensurando devemos aplicar um fator de correção, aproximando a distribuição de pequenos valores a uma distribuição normal.

Este fator, conhecido como fator de abrangência (k), é tabelado em função do tamanho da amostra **n** (ou do grau de liberdade **ν = n-1**) e da probabilidade de abrangência desejada **p**.



Como pudemos observar, quanto maior o número de amostras, e consequentemente o número de graus de liberdade, mais a curva t-Student se aproxima da curva normal.

Na Metrologia, seguindo as diretrizes do documento internacional Guia para Expressão da Incerteza de Medição – GUM, devemos considerar a probabilidade de abrangência de 95,45%.

Se não tivermos a média da população μ (como na distribuição normal) e sim a média da amostra $\bar{x}$ (distribuição t-Student), verificaremos que a média da população sempre estará compreendida no intervalo $\bar{x} \pm k.(s_{\bar{x}})$, onde **k** é o fator de abrangência e $(s_{\bar{x}})$ é o desvio padrão da média.

A tabela a seguir apresenta o fator **k** em função do tamanho da amostra **n** (ou do grau de liberdade **ν = n-1**) e da probabilidade de abrangência desejada **p**.

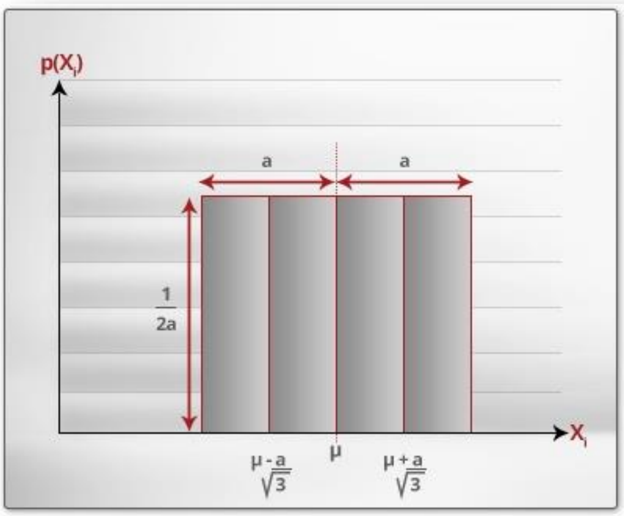
Fator de Abrangência
Ou seja, x somado ou subtraído do resultado da multiplicação de k por s.

n	ν (n-1)	PROBABILIDADE DE ABRANGÊNCIA					
		68,27%	90%	95%	95,45%	99%	99,73%
2	1	1,84	6,31	12,71	13,97	63,66	235,8
3	2	1,32	2,92	4,30	4,53	9,92	19,21
4	3	1,20	2,35	3,18	3,31	5,84	9,22
5	4	1,14	2,13	2,78	2,87	4,60	6,62
6	5	1,11	2,02	2,57	2,65	4,03	5,51
7	6	1,09	1,94	2,45	2,52	3,71	4,90
8	7	1,08	1,89	2,36	2,43	3,50	4,53
9	8	1,07	1,86	2,31	2,37	3,36	4,28
10	9	1,06	1,83	2,26	2,32	3,25	4,09
11	10	1,05	1,81	2,23	2,28	3,17	3,96
12	11	1,05	1,80	2,20	2,25	3,11	3,85
13	12	1,04	1,77	2,16	2,21	2,98	3,69
14	13	1,04	1,77	2,16	2,21	3,01	3,69
15	14	1,04	1,76	2,14	2,20	2,98	3,64
16	15	1,03	1,75	2,13	2,18	2,95	3,59

17	16	1,03	1,75	2,12	2,17	2,92	3,54
18	17	1,03	1,74	2,11	2,16	2,90	3,51
19	18	1,03	1,73	2,10	2,15	2,88	3,48
20	19	1,03	1,73	2,09	2,14	2,86	3,45
21	20	1,03	1,72	2,09	2,13	2,85	3,42
22	21	1,02	1,72	2,08	2,13	2,83	3,40
23	22	1,02	1,72	2,07	2,12	2,82	3,38
24	23	1,02	1,71	2,07	2,11	2,81	3,36
25	24	1,02	1,71	2,06	2,11	2,80	3,34
26	25	1,02	1,71	2,06	2,11	2,79	3,33
27	26	1,02	1,71	2,06	2,10	2,78	3,32
28	27	1,02	1,70	2,05	2,10	2,77	3,30
29	28	1,02	1,70	2,05	2,09	2,76	3,29
30	29	1,02	1,70	2,05	2,09	2,76	3,28
∝	∝	1,00	1,64	1,96	2,00	2,58	3,00

7. Distribuição retangular ou uniforme

Quando a distribuição de probabilidade for constante num intervalo definido, estaremos diante de uma distribuição uniforme ou retangular. Perceba que, dentro do intervalo em torno do valor central μ, a probabilidade de ocorrência de que qualquer um desses valores é rigorosamente a mesma.



Onde o desvio padrão da distribuição retangular é dado pela equação abaixo:

$$s = \frac{a}{\sqrt{3}}$$

Onde **a** equivale à metade do intervalo atribuído à distribuição.

Exemplo:

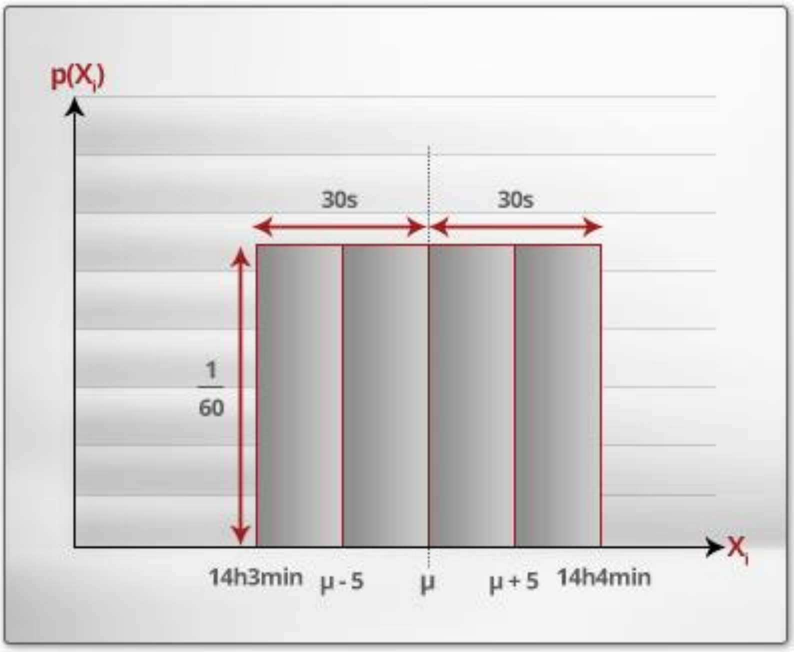
Suponha que você esteja olhando para um relógio digital que marca 14:03.

Seguindo as regras estabelecidas pelo Sistema Internacional de Unidades devemos expressar essa medida do tempo da seguinte forma: 14h 3min.

Agora imagine que você queira saber até mesmo o valor dos segundos.

Como o relógio não marca os segundos existe certa incerteza acerca do intervalo de tempo que vai de 14h 3min até 14h 4min. Certo?

Como nesse intervalo a probabilidade dos valores é igual teremos, então, a seguinte distribuição retangular:



Onde *a* é igual a 30 segundos. Logo, o desvio padrão do período de um minuto é igual a:

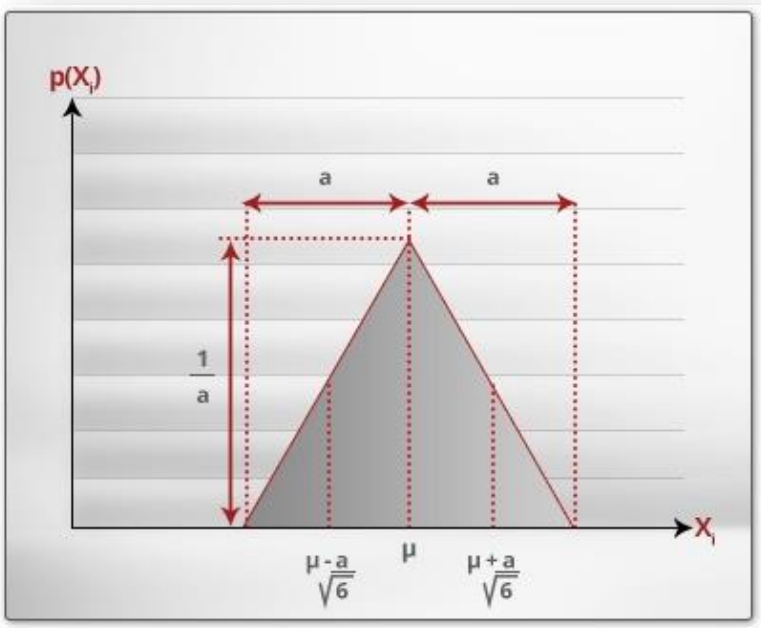
$$s = \frac{30}{\sqrt{3}} = 17{,}2\ segundos$$

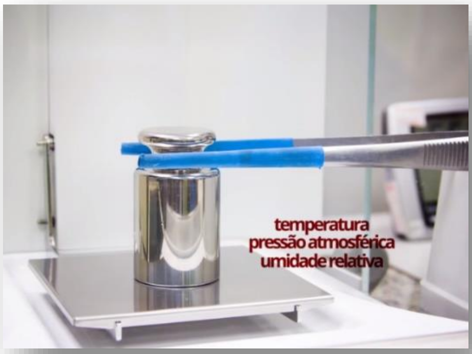
8. Distribuição triangular

Quando a distribuição de probabilidade for maior na parte central, num intervalo definido, e decair linearmente nas extremidades, estaremos diante de uma distribuição triangular.

Onde o desvio padrão da distribuição triangular é dado pela equação abaixo:

$$s = \frac{a}{\sqrt{6}}$$



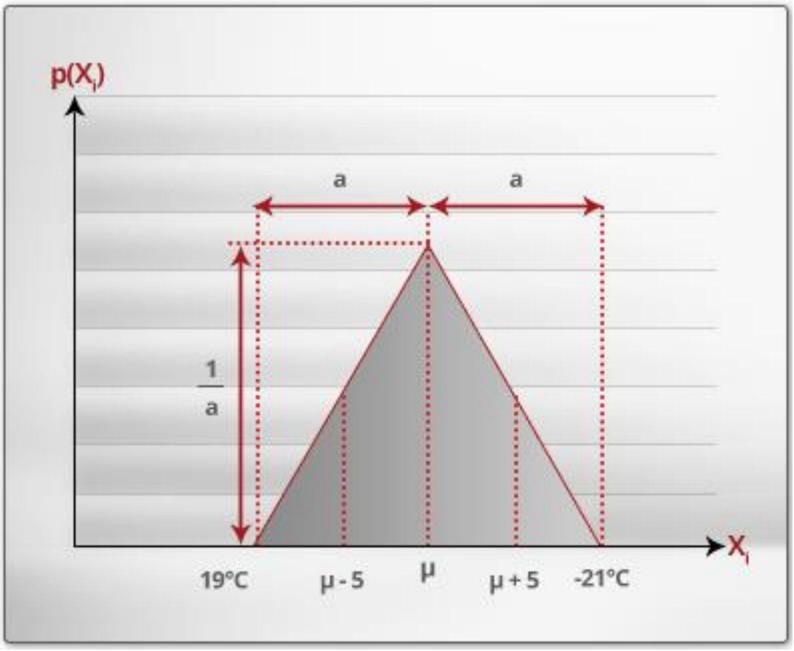


Exemplo:

Na calibração de uma balança você percebeu que a temperatura do laboratório variou de 19ºC a 21ºC.

Como estamos tratando de variação de condições ambientais, na qual há um valor central mais provável e ainda desconhecido, e dois valores extremos com menor probabilidade de ocorrência (19ºC e 21ºC), essas informações sugerem uma distribuição de probabilidade triangular.

Logo teríamos:



Onde "a" é igual a 1ºC. Logo, o desvio padrão da variação da temperatura é igual a:

$$s = \frac{1}{\sqrt{6}} = 0{,}4°C$$

Bom, a aula de hoje fica por aqui…

Na próxima aula você verá que erro e incerteza não são a mesma coisa e, além disso, existem tipos diferentes de erros e de incerteza. Não fique fora dessa!

Lembre-se de fazer os exercícios de fixação e se ficar com alguma "incerteza", acesse o fórum de dúvidas da aula dois que teremos prazer em lhe ajudar!

Bons estudos!