Introdução a tópicos de estatística em FAE (Parte 2)

Sandro Fonseca de Souza

Sumário

- Motivações
    - Como os estatísticos e físicos colaboram?
- O que nós queremos saber através da Estatística em Física de Altas Energias (FAE)
- Os Experimentos em FAE
    - Quais o passos de uma análise em FAE?
- Análise estatística
    - Origens históricas
    - Métodos estatísticos
    - Testes de significância e de hipóteses
        - Testes unilaterais
            - p-value e níveis de significância
        - Testes bilaterais
            - nível de confiança
            - Intervalo de confiança
        - Significância e descoberta em FAE

Motivações

A Física de Partículas surgiu como uma disciplina independente há meio século.

Foi pioneira na “big science"; experimentos são realizados em aceleradores de energia e complexidade crescentes por colaborações de muitos físicos de muitos institutos em diversas parte do mundo.

Desenvolveu uma metodologia de pesquisa na qual a estatística é de grande importância e nos últimos anos tem trabalhado fortemente com os especialistas em estatística solidificando as técnicas e métodos usados em outras áreas do conhecimento.

Os estatísticos estão familiarizados com os métodos de pesquisa e questões estatísticas que surgem em, digamos, testes de campo agrícolas ou testes clínicos.

PhyStat Workshop

Collaboration Workspaces

BROWSE PAGE PUBLISH FOLL

Search this site

PHYSTAT Workshop Homepage

PhyStat Workshop Homepage [Link para página]

- Home
- Notebook
- Documents
- Recent
- Site Contents

Introduction:

PHYSTAT is a workshop series dealing with statistical methods in particle physics. It was founded in 2000 by Louis Lyons. A wide range of topic is covered: frequentist and bayesian inference, parameter estimation, hypothesis testing, Goodness of Fit testing, confidence interval estimation, unfolding, multivariate analysis techniques, systematic uncertainties, data combination and more. The workshop is a unique meeting point of Physicists and Statisticians, where the latest advances in statistical techniques and procedures are exchanged. The workshop consists of a mixture of invited and contributed talks plus panel discussions. The proceedings of the workshop provide a comprehensive reference on statistical issues and state of the art methods used in particle physics and neighbouring fields.

Since beginning of 2018 Olaf Behnke (DESY) is actaing as chair person of PHYSTAT.

Upcoming Events:

- 7th July 2021 Xiao-Li Meng (Harvard) "From COVID-19 testing to election prediction: how small are our big data? "
- we are currently preparing further seminars to take place later in 2021

Please send your suggestions for possible future meetings or seminars to olaf.behnke@desy.de,

Links to past events: PHYSTAT seminars joint with CERN EP-IT Data Science:

- 12th May 2021 Richard Samworth (Cambridge) "USP: an independence test that improves on Pearson's chi-squared and the G-test "
- 14th Apr 2021 Chad Shafer (CMU): "An Overview of Nonparametric Regression"
- 10th Mar 2021 Nick Wardle (Imperial College): "Can we really reinterpret data from the LHC?"
- 20th Jan 2021 Aaditya Ramdas (CMU): "Some new concepts in methodological statistics, that could be of utility in the sciences"
- 9th Dec 2020 Roberto Trotta (Imperial College) : "Supervised learning with biased Training data and applications to Supernovae type 1a Cosmology
- 11th Nov 2020 Wolfgang Rolke (UPR): "Testing Goodness of Fit"
- 28th Oct 2020 Larry Wasserman (CMU): "Optimal Transport"
- 14th Oct 2020 Kyle Cranmer (NYU) : "Likelihood publishing, RECAST, and simulation based inference"
- 15th July 2020 Chad Shafer (CMU): "An Overview of Approxumate Bayesian Computation"
- 3rd June 2020 Harrison Prosper (FSU): "Statistics in Astronomy: A View Through the Looking Glass"
- 1th April 2020 Mikael Kuusela (CMU): "Uncertainty Quantification in ill-posed inverse problems: Case Studies in the Physical Sciences "
- 5th Feb 2020 at CERN Yoav Benjamini (Tel Aviv) "Addressing the effect of selection on inference with the False Discovery Rate"
- 15th Jan 2020 at CERN Allen Caldwell (MPIM): "Accelerating Bayesian Computation: Parallelizing Markov Chain MC"
- 20th Nov 2019 at CERN Christian Mueller (LMU): "A general introduction to continuous optimization"

https://cerncourier.com/a/the-history-and-future-of-the-phystat-series/

By physicists, for physicists

G. Cowan, Statistical Data Analysis, Clarendon Press, Oxford, 1998.

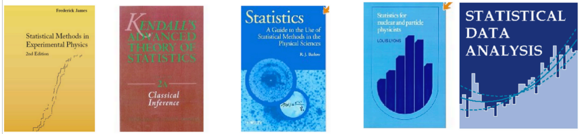
R.J.Barlow, A Guide to the Use of Statistical Methods in the Physical Sciences, John Wiley, 1989;

F. James, Statistical Methods in Experimental Physics, 2nd ed., World Scientific, 2006;

- W.T. Eadie et al., North-Holland, 1971 (1st ed., hard to find);

S.Brandt, Statistical and Computational Methods in Data Analysis, Springer, New York, 1998.

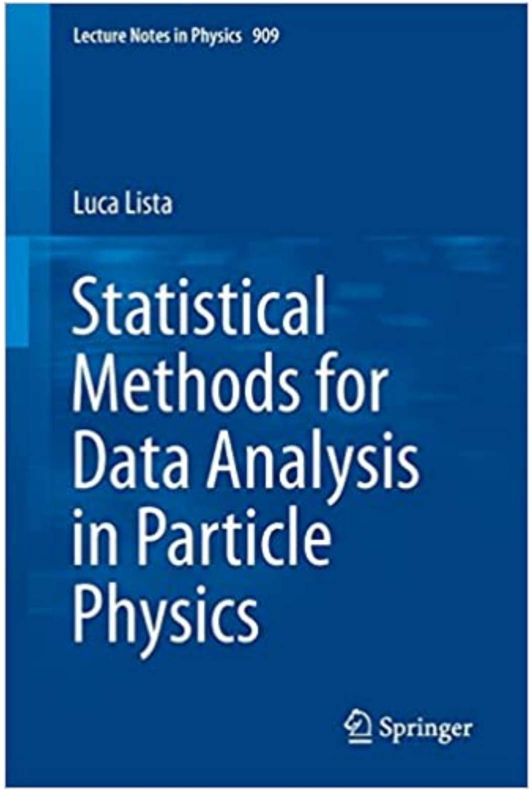
L.Lyons, Statistics for Nuclear and Particle Physics, CUP, 1986.



My favorite statistics book by a statistician:

Stuart, Ord, Arnold. "Kendall's Advanced Theory of Statistics" Vol. 2A *Classical Inference & the Linear Model.*

Mais algumas referências



O que nós queremos saber através da Estatística em FAE

- Algumas questões que queremos responder:
  - Existem partículas supersimétricas?
  - Porque ainda estamos testando propriedades do bóson de Higgs?
  - Qual a seção de choque?
  - Qual a massa do bóson?

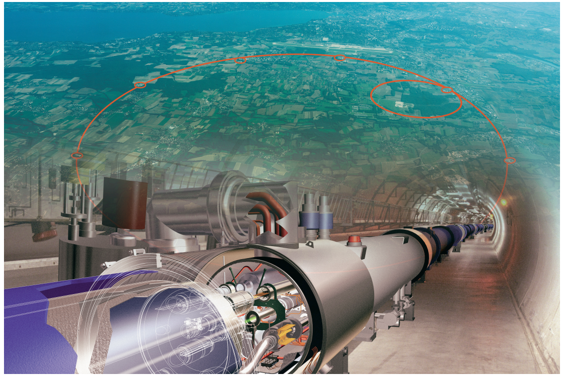
- Os testes estatísticos constroem modelos probabilísticos para o processo de interesse:
  - Teste de hipotese
  - Intervalos de confiança (medidas & incertezas)

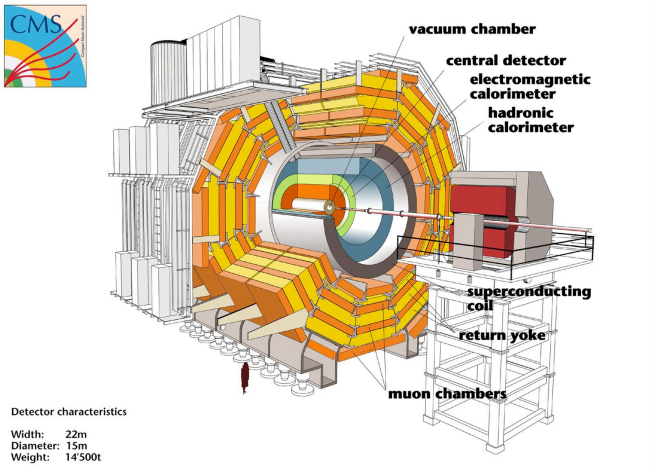
- Utilizamos testes estatísticos para obter os resultados finais

https://atlas.cern/updates/feature/higgs-boson https://videos.cern.ch/record/2757407

Experimentos em FAE

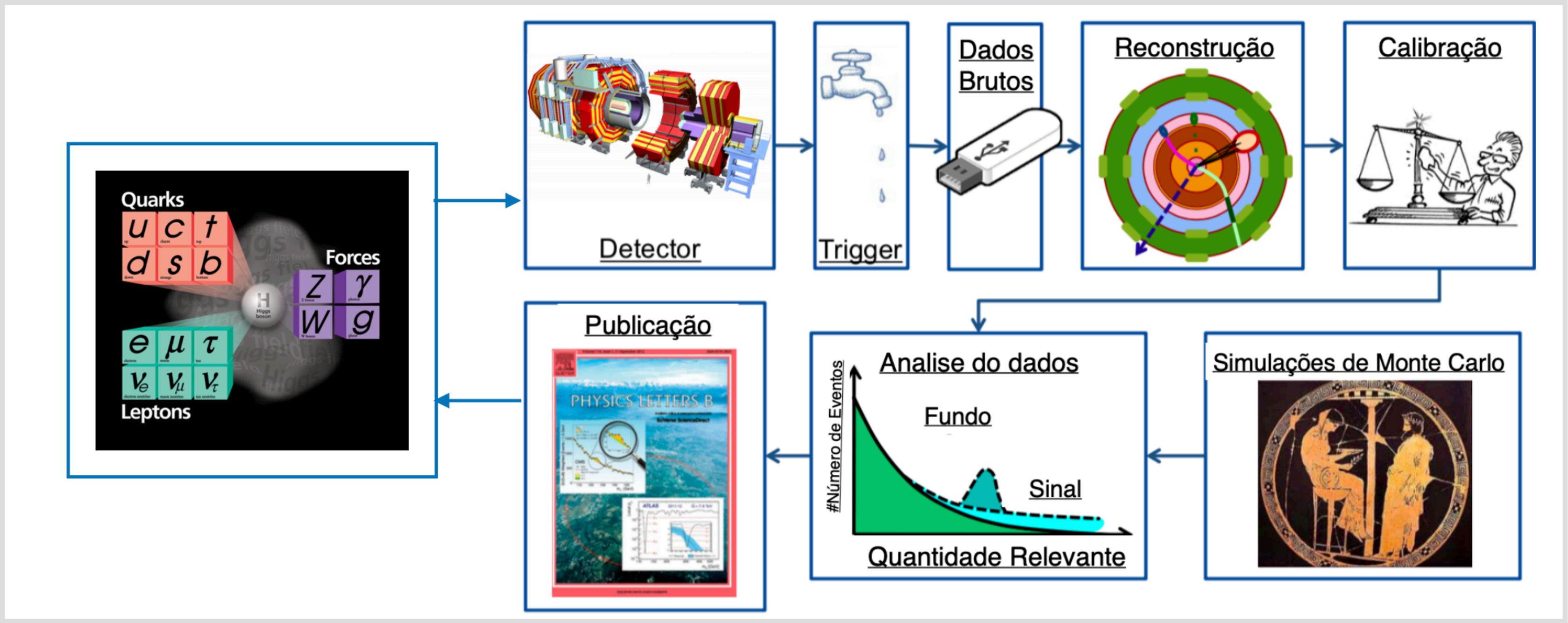
- As interações fundamentais da natureza se manifestam em colisões de partículas em altas energias.
- Os aceleradores de partículas proporcionam que essas interações ocorram, continuamente, em uma pequena região do espaço (vértice da colisão).
- A partir desses vértices outras partículas secundárias são criadas (eventos), e se afastam em todas as direções, em um amplo intervalo de energia.





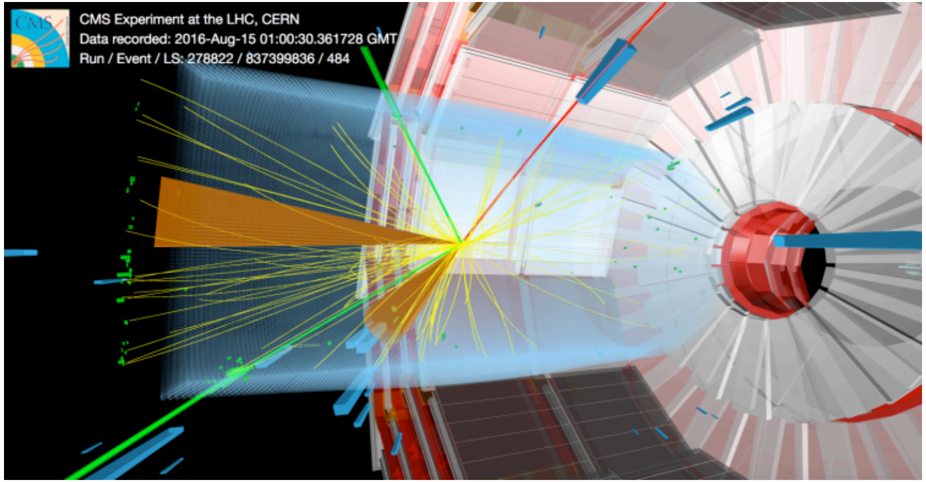
Fonte: Notas do Prof. Vitor Oguri

Jornada em FAE



Objetivos gerais

- Identificação de partículas (elétrons, múons, neutrinos, quarks, bósons mediadores, hádrons)
- Determinação de atributos intrínsecos (carga, massa, momento magnético, spin, tempo de vida, razão de ramificação)
- Determinação de parâmetros associados aos fenômenos envolvidos (seção de choque, constantes de acoplamento)

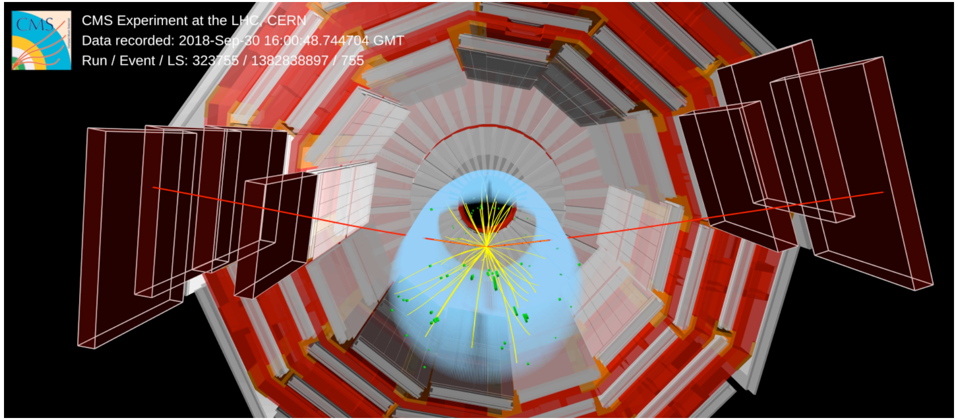


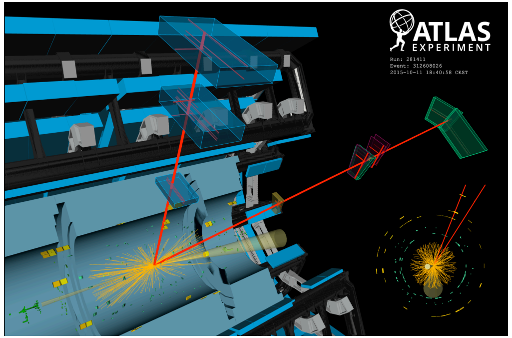
https://cms.cern/news/cms-collaboration-has-observed-new-production-mechanism-top-quarks

Fonte: Notas do Prof. Vitor Oguri

Características mensuráveis

- Posições, trajetórias, velocidades e tempos de vôo
- Distribuições de grandezas geométricas (ângulos polar e azimutal, pseudo-rapidity) e cinemáticas (momentum, energia)





https://news.fnal.gov/2020/08/cern-experiments-announce-first-indications-of-a-rare-higgs-boson-process/

Fonte: Notas do Prof. Vitor Oguri

Jornada em FAE na prática

Dados simulados utilizando a técnica de MC (os processos físicos de interesse [sinal e fundos] e com as contribuições da simulação do detector)

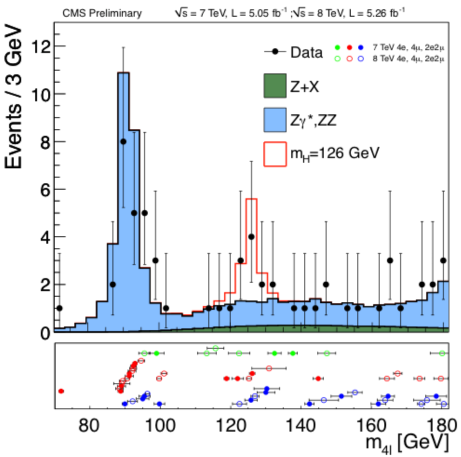
Dados reais disponibilizados pelo Experimento

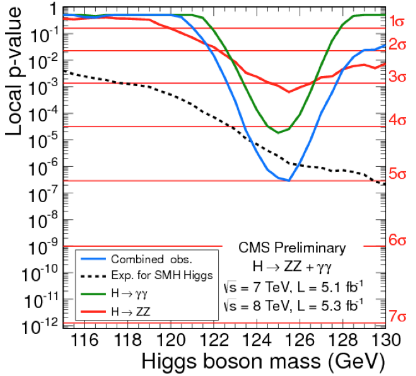
Análise Estatística

Aplicar as estratégias de seleção de eventos
- n-tuples
- cut flows - seleção de variáveis
- técnicas de análise multivariável (redes neurais artificiais ou algoritmos de aprendizagem de máquina, redes bayesianas)

Medida de interesse

Descoberta





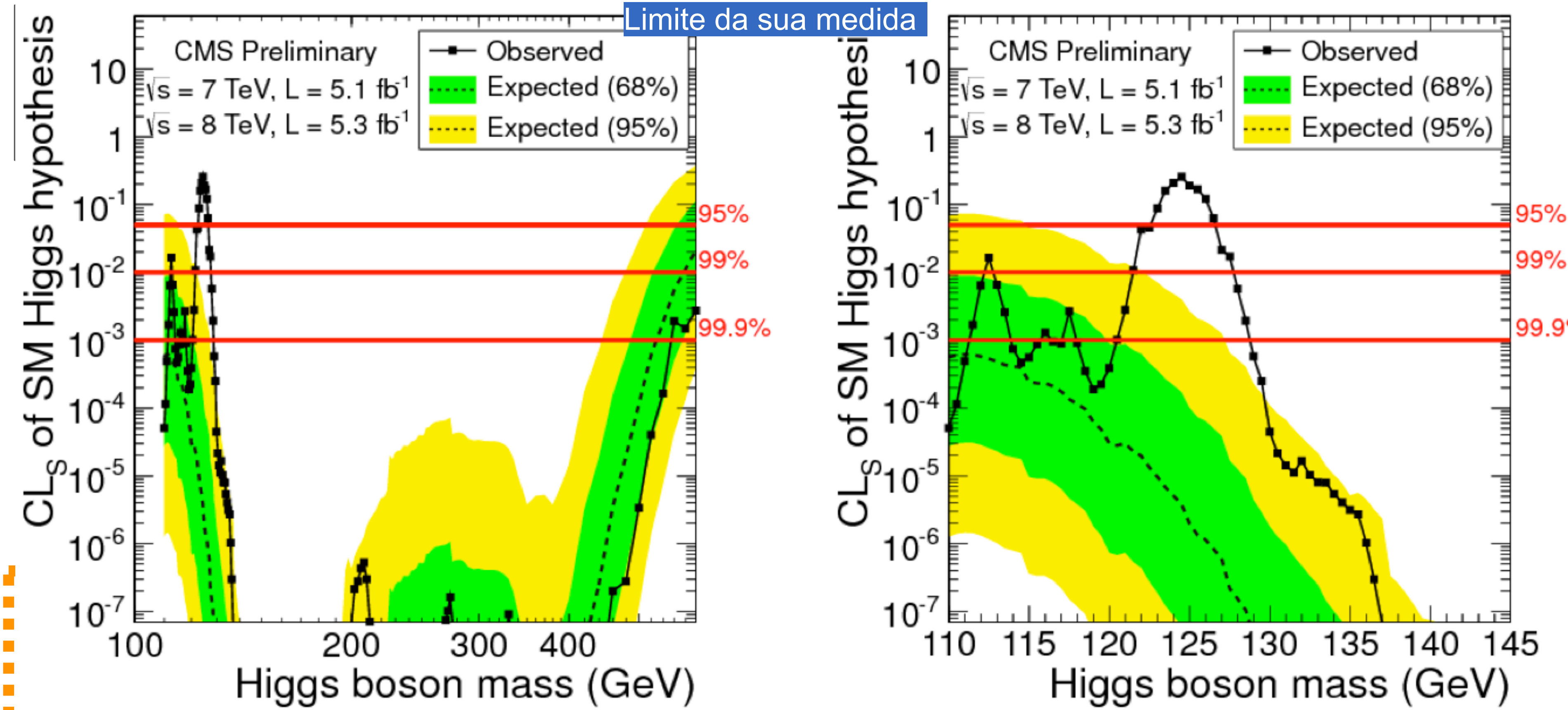
Análise Estatística

Seleção final dos seus eventos de MC

Seleção final dos seus eventos reais de interesse

Jornada em FAE na prática

Limite da sua medida



interesse

Análise estatística e as suas origens

Karl Pearson (1857-1936)

- The Grammar of Science – 1892
  - Modelos estatísticos como alternativas à visão determinística do séc. XIX
  - Todo experimento está sujeito a efeitos imprevistos e não observáveis
  - Os resultados $\{x_i\}$ de um experimento obedecem a certas distribuições estatísticas, $\rho(x)$, que são caracterizadas por alguns parâmetros: média, desvio-padrão, simetria e curtose[1]
  - Os observáveis na Ciência são as funções de distribuições estatísticas que descrevem as probabilidades associadas às observações
  - Método dos momentos e teste de χ^2 (1900)

https://pt.wikipedia.org/wiki/Curtose

Fonte: Notas do Prof. Vitor Oguri

Statistical Methods for Research Workers – 1925 The Design of Experiments – 1935

Ronald Fisher (1890-1962)

- Todo experimento deve começar com um modelo matemático que estime os resultados esperados
- Os estimadores são aleatórios e devem ser avaliados segundo suas distribuições de probabilidades, e aos seguintes critérios:
  - consistência – quanto maior o tamanho (N) de uma amostra mais próximo um estimador a' estará do valor esperado $\langle a \rangle$ do parâmetro $\lim a' = \langle a \rangle$ $(N \to \infty)$
  - eficiência – quanto menor a variância associada mais eficiente é o estimador $V(a_1) > V(a_2) \Rightarrow a_2$ mais eficiente
  - não-tendenciosidade – o valor médio de um estimador deve tender ao valor esperado do parâmetro $a' \to \langle a \rangle$
- Graus de liberdade (1922) e Likelihood (1925)

Fonte: Notas do Prof. Vitor Oguri

Pearson x Fisher

- Pearson
  - As distribuições descrevem verdadeiramente os dados (medidas) resultantes de um experimento, no sentido de que a partir de um grande número de medições pode-se determinar os parâmetros da distribuição real das medidas.
- Fisher
  - As medições resultam de uma população hipotética, e a partir de um experimento obtém-se apenas os estimadores dos parâmetros das distribuições hipotéticas dos dados.

Fonte: Notas do Prof. Vitor Oguri

Método da Máxima Verossimilhança de Fisher

- $\rho(x_i|\theta) \rightarrow$ distribuição de probabilidades (PDF) para as medidas de uma grandeza x, onde θ é um parâmetro da distribuição
- para uma amostra A = $\{x_1, x_2, x_3 \ldots x_N\}$ de N medidas independentes, a probabilidade associada é dada por N

$$\prod_{i=1}^{N} \rho(x_i|\theta)$$

- se a PDF ρ e o parâmetro são corretos, espera-se que em relação a qualquer outra distribuição ou parâmetro, essa probabilidade seja máxima.

Fonte: Notas do Prof. Vitor Oguri

Método da Máxima Verossimilhança de Fisher

- A função definida por:

$$\mathcal{L}(\theta|x_1, x_2, x_3 \ldots x_N) = K \prod_{i=1}^{N} \rho(x_i|\theta)$$

quantifica o quão verossímil é o valor do parâmetro para uma dada amostra de dados, no sentido que, se θ_A e θ_B são dois valores possíveis para o parâmetro e

$$\mathcal{L}(\theta_A) > \mathcal{L}(\theta_B)$$

diz-se que θ_A é uma estimativa mais verossímil para o parâmetro do que θ_B.

Fonte: Notas do Prof. Vitor Oguri

Método da Máxima Verossimilhança de Fisher

- A função definida por:

$$\mathcal{L}(\theta|x_1, x_2, x_3 \dots x_N) = K \prod^{N} \rho(x_i|\theta)$$

- Os estimadores de máxima verossimilhança (ML) são aqueles que maximizam a chamada função de verossimilhança (likelihood).
- Fisher mostrou que esses estimadores são "geralmente" consistente, os mais eficientes, e a tendenciosidade pode ser calculada e, portanto, subtraída.

diz-se que θ_A é uma estimativa mais verossímil para o parâmetro do que θ_B.

Fonte: Notas do Prof. Vitor Oguri

Exercício bônus 1

- De uma caixa que contém 10 bolas, entre vermelhas e azuis, é extraída uma amostra com reposição de 3 bolas vermelhas e uma azul. Quantas bolas azuis há na caixa?
- O estudante deve realizar a parte analítica e reproduzir os histogramas propostos com as respectivas PDFs.
- Referencias páginas 95-97 do livro do Vitor Oguri.

Fonte: Notas do Prof. Vitor Oguri

Estimativas de Parâmetros

- Limite da função de verossimilhança

  - espera-se que a função de verossimilhança, como função de um parâmetro θ, represente uma distribuição de “probabilidades” cujo máximo seja dado por um valor situado em torno do valor esperado para o parâmetro.
  - para um número “suficientemente grande” de medidas, pode-se expandir o logaritmo da função de verossimilhança em torno de seu máximo,

$$\ln \mathcal{L}(\theta) = \ln \mathcal{L}(\hat{\theta}) - \frac{1}{2}\left|\frac{\partial^2 \ln \mathcal{L}}{\partial \theta^2}\right|_{\hat{\theta}} (\theta - \hat{\theta})^2$$

Fonte: Notas do Prof. Vitor Oguri

Limite da função de verossimilhança

- para grandes amostras (teorema do limite central), a função de verossimilhança pode ser aproximada por uma distribuição gaussiana

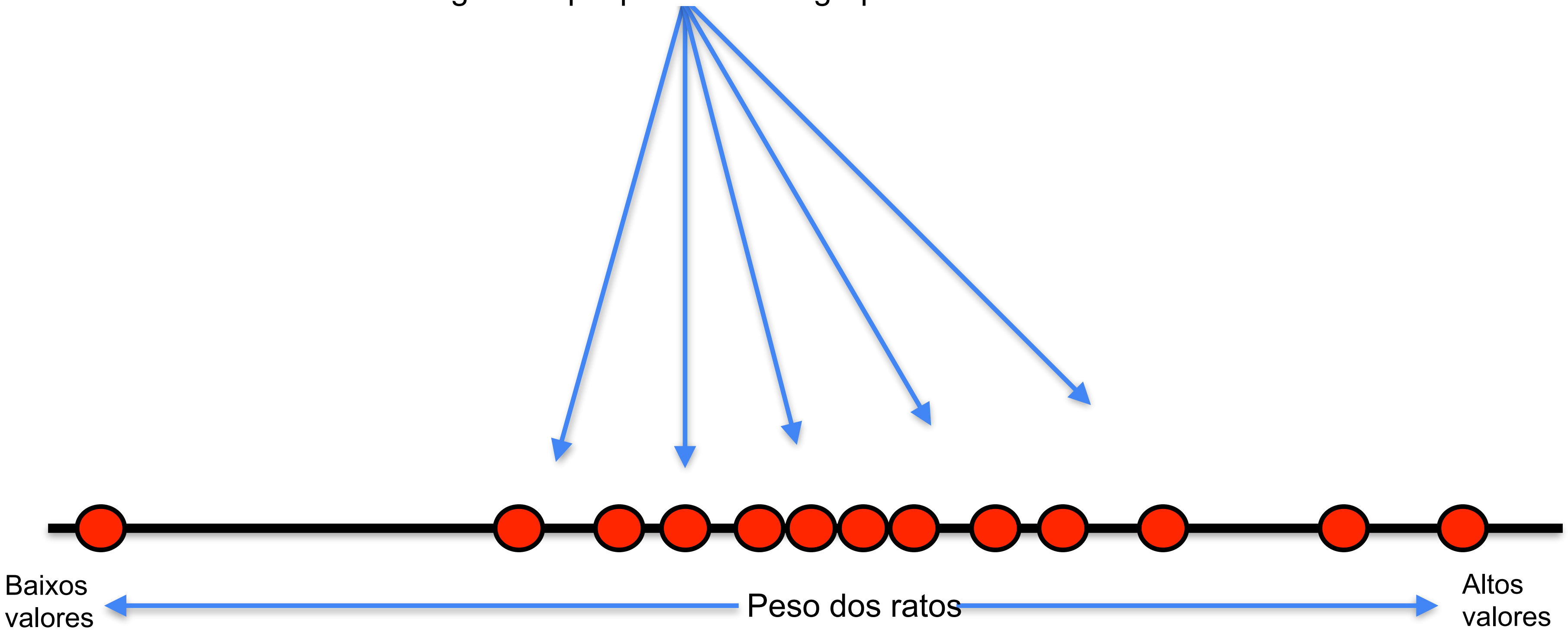
$$\boxed{\mathcal{L}(\theta) \propto e^{-\frac{1}{2}\left|\frac{\partial^2 \ln \mathcal{L}}{\partial \theta^2}\right|_{\hat{\theta}}(\theta - \hat{\theta})^2}}$$

Fonte: Notas do Prof. Vitor Oguri

https://towardsdatascience.com/probability-concepts-explained-maximum-likelihood-estimation-c7b4342fdbb1

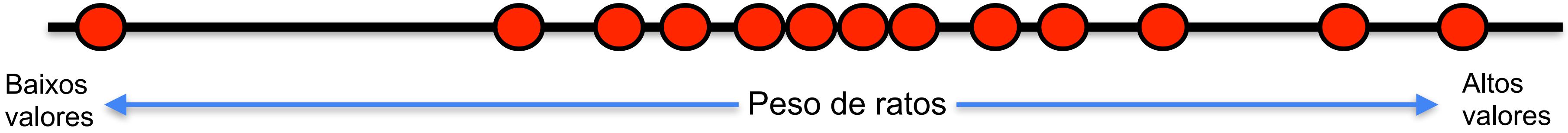
Método da Máxima Verossimilhança de Fisher

Digamos que pesamos um grupo de ratos



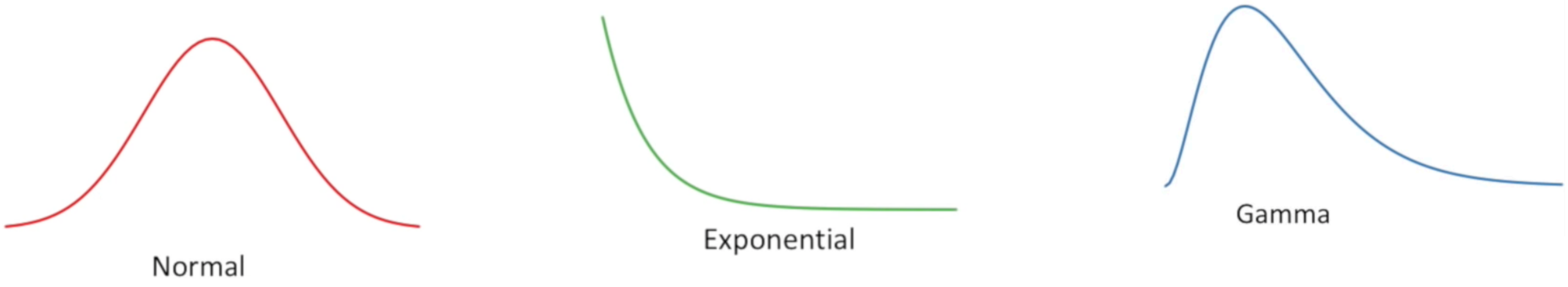
Método da Máxima Verossimilhança de Fisher

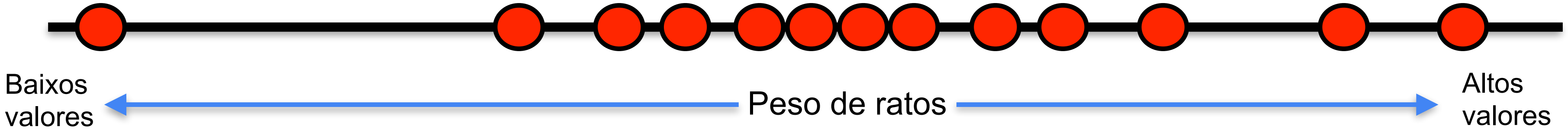
O objetivo da técnica de máxima verossimilhança é encontrar uma forma otimizada para ajustar uma distribuição para os dados.



Método da Máxima Verossimilhança de Fisher

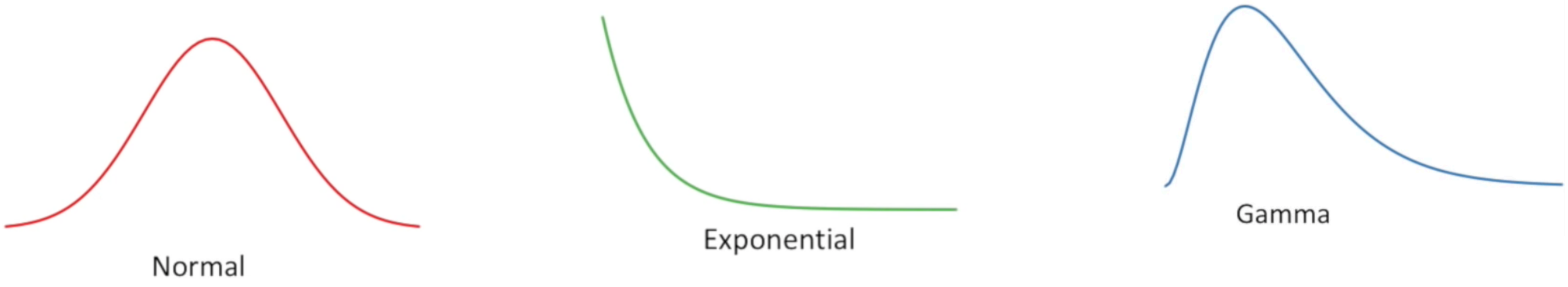
Existem diferentes tipos de distribuições para diferente tipos de dados

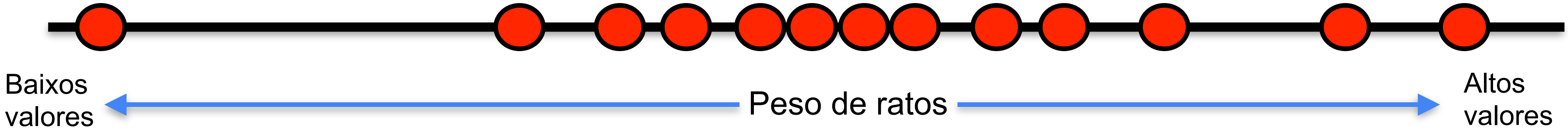




Método da Máxima Verossimilhança de Fisher

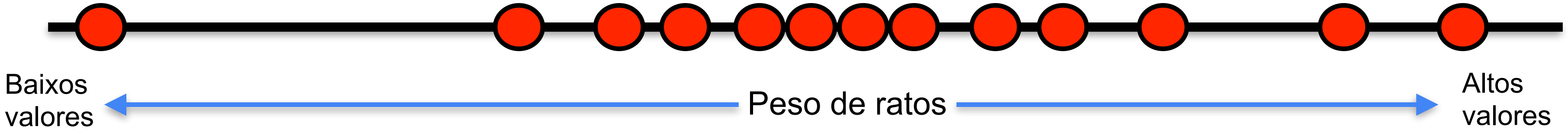
A razão que você quer ajustar a distribuição para seus dados é que isso pode ser fácil trabalhar com isso e isso é também mais geral. Isso aplica para cada experimento de um mesmo tipo.





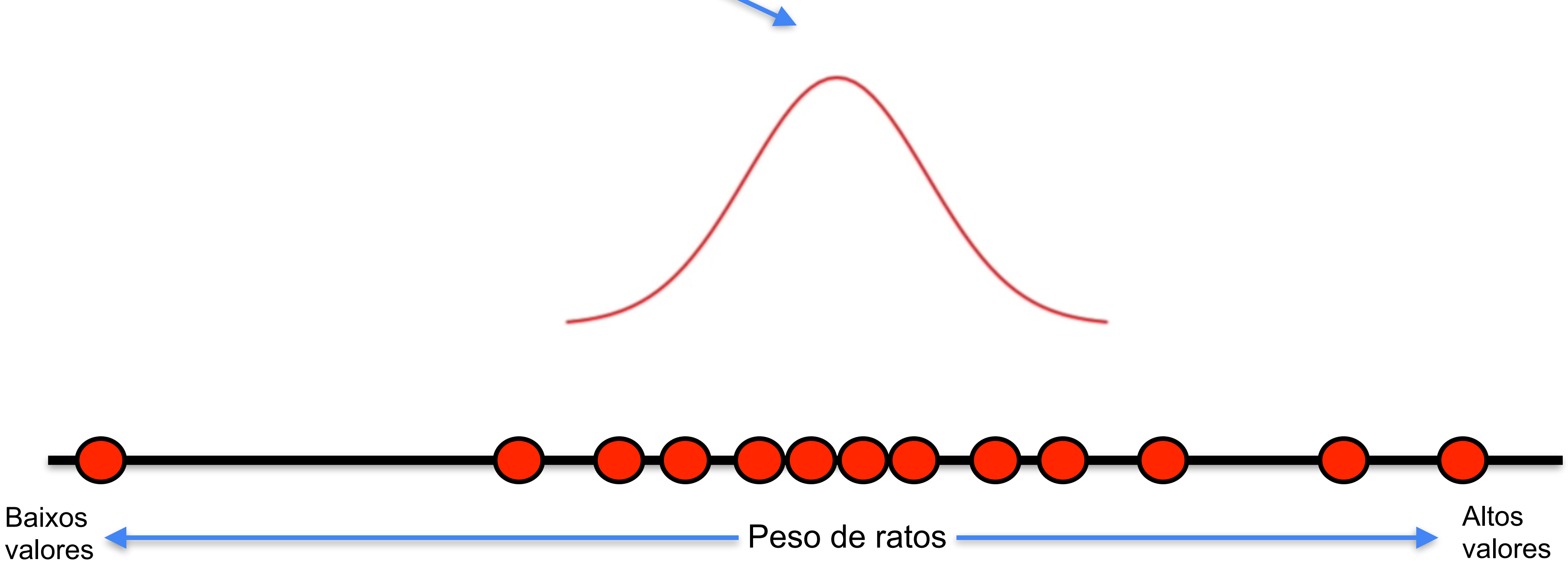
Método da Máxima Verossimilhança de Fisher

Nós pensamos que os pesos devem ser distribuídos normalmente…



Método da Máxima Verossimilhança de Fisher

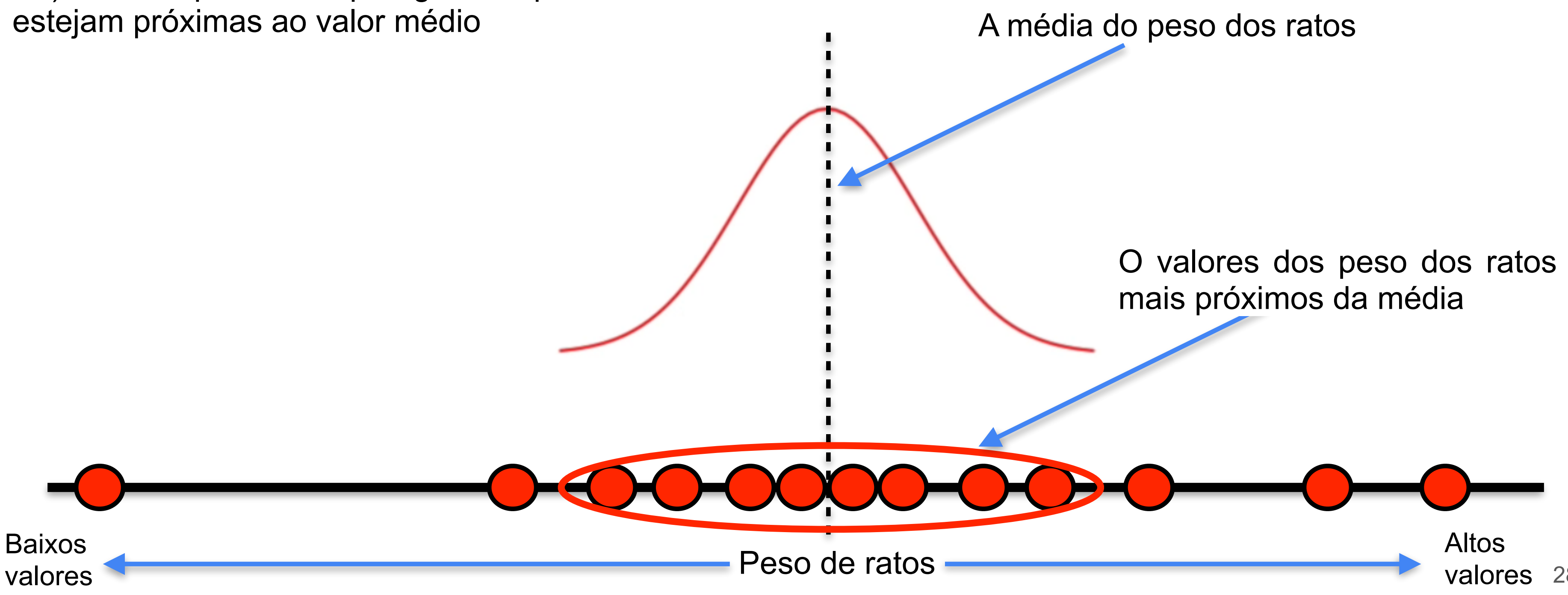
Isso significa que achamos que veio desse tipo de distribuição



Método da Máxima Verossimilhança de Fisher

"Normalmente distribuído" significa um monte de coisas

1) Nós esperamos que grande parte das medidas estejam próximas ao valor médio

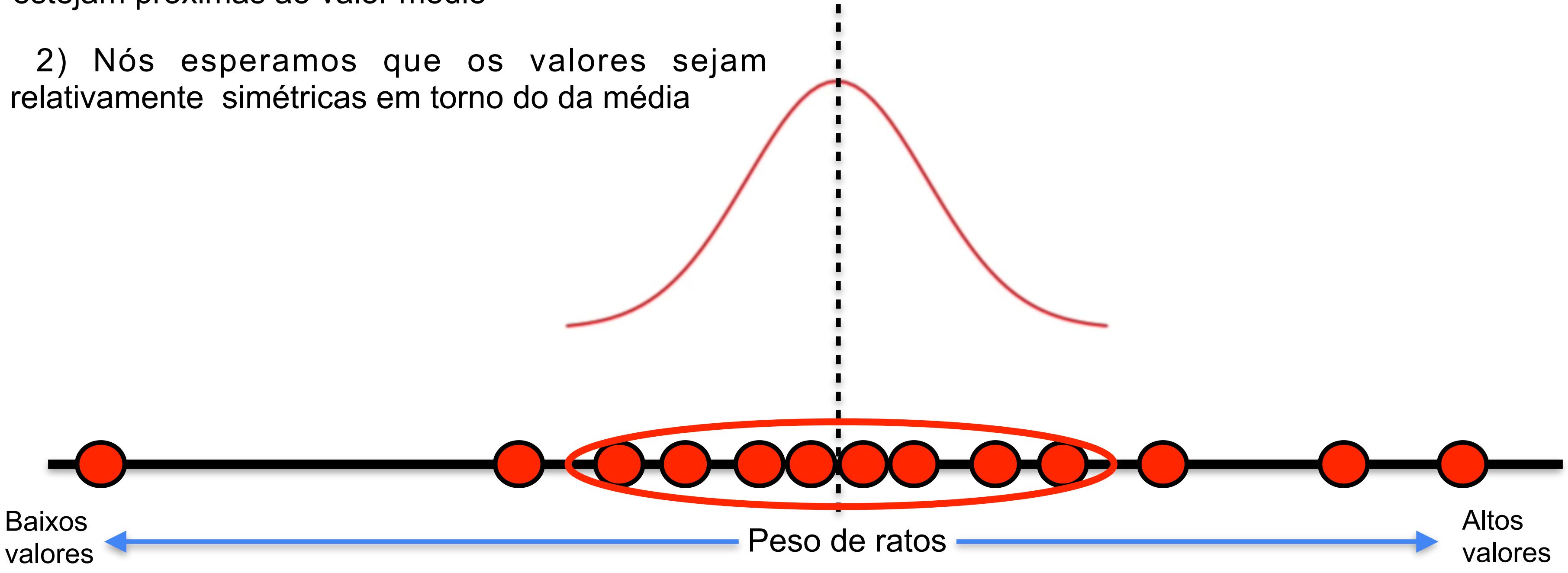


Método da Máxima Verossimilhança de Fisher

"Normalmente distribuído" significa um monte de coisas

1) Nós esperamos que grande parte das medidas estejam próximas ao valor médio

2) Nós esperamos que os valores sejam relativamente simétricas em torno do da média



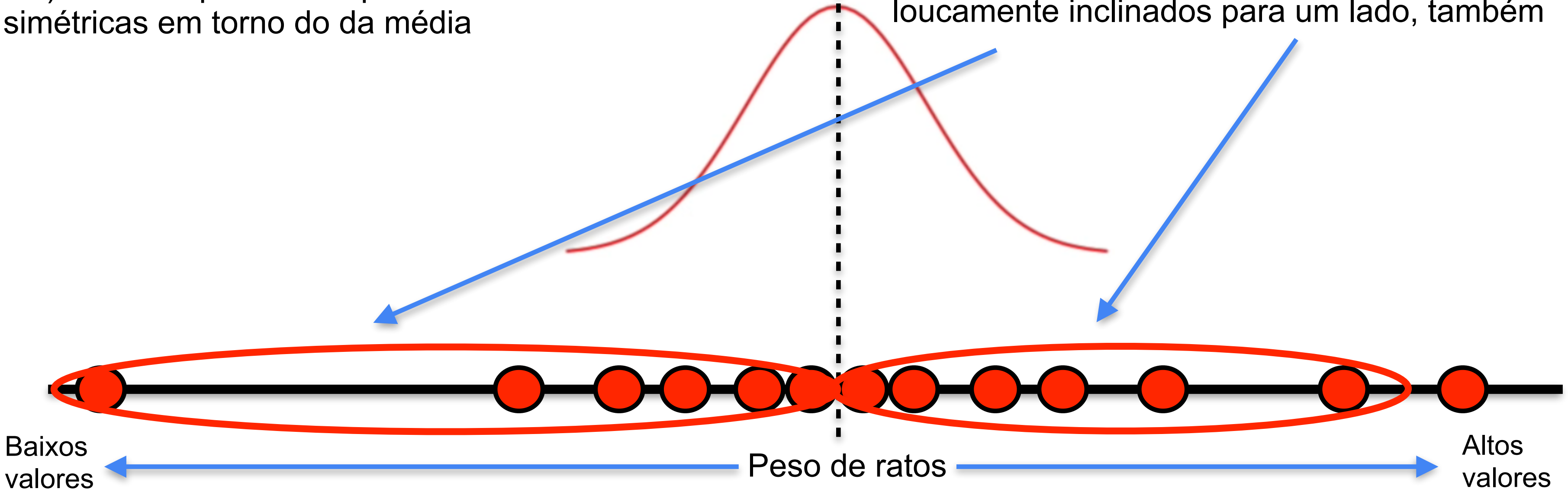
Método da Máxima Verossimilhança de Fisher

"Normalmente distribuído" significa um monte de coisas

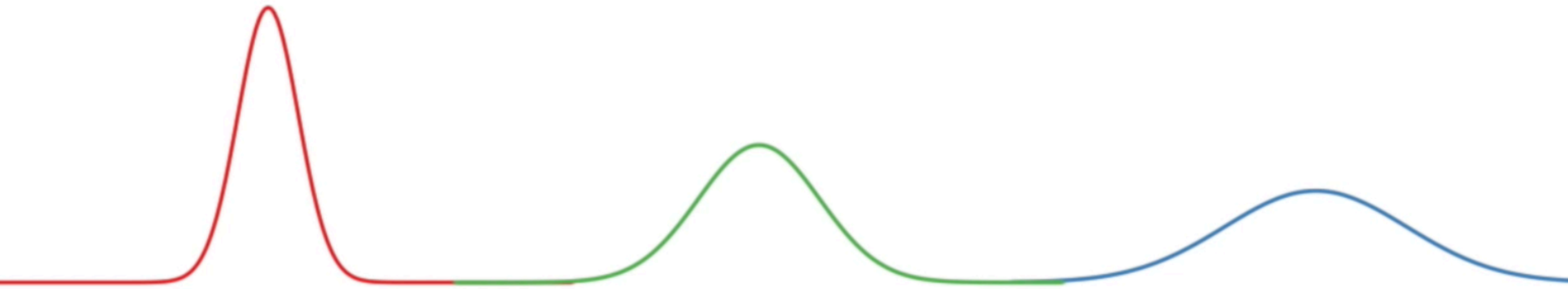
1) Nós esperamos que grande parte das medidas estejam próximas ao valor médio

2) Nós esperamos que as são relativamente simétricas em torno do da média

Apesar das medidas não serem perfeitamente simétricas em torno da média, elas não são loucamente inclinados para um lado, também



Existem distribuições normais de todos os tipos e tamanhos

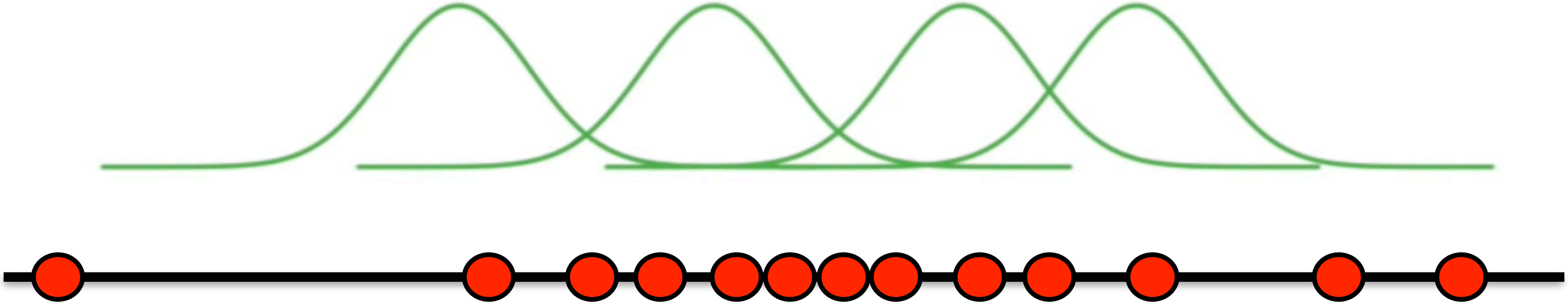


"Estreita"

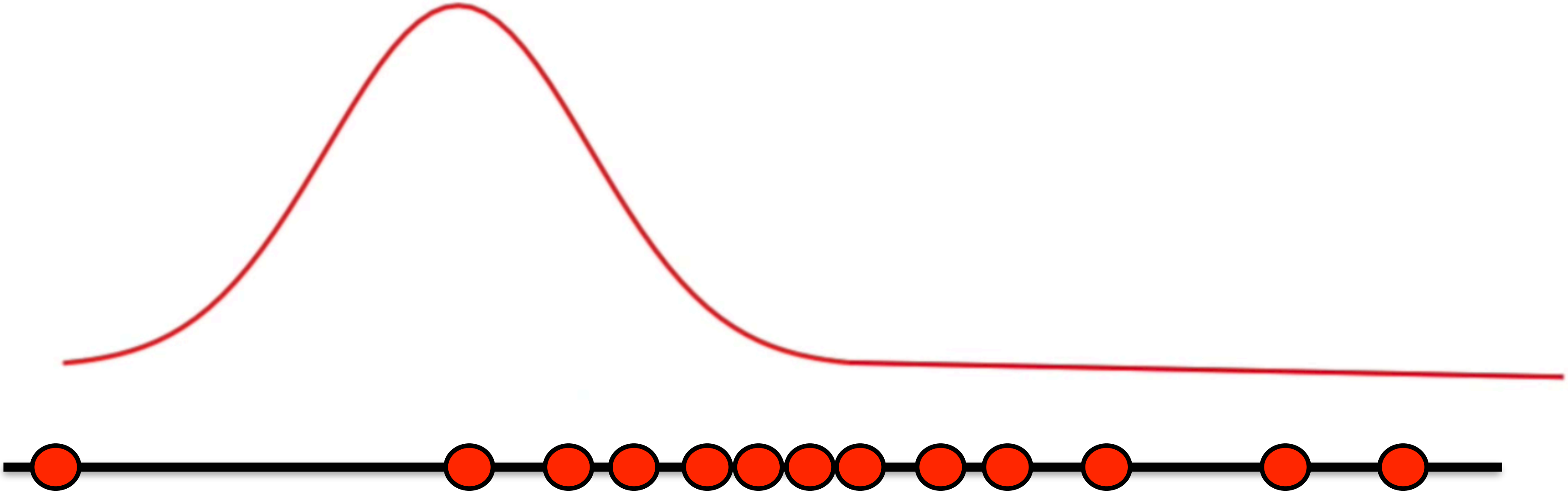
"Média"

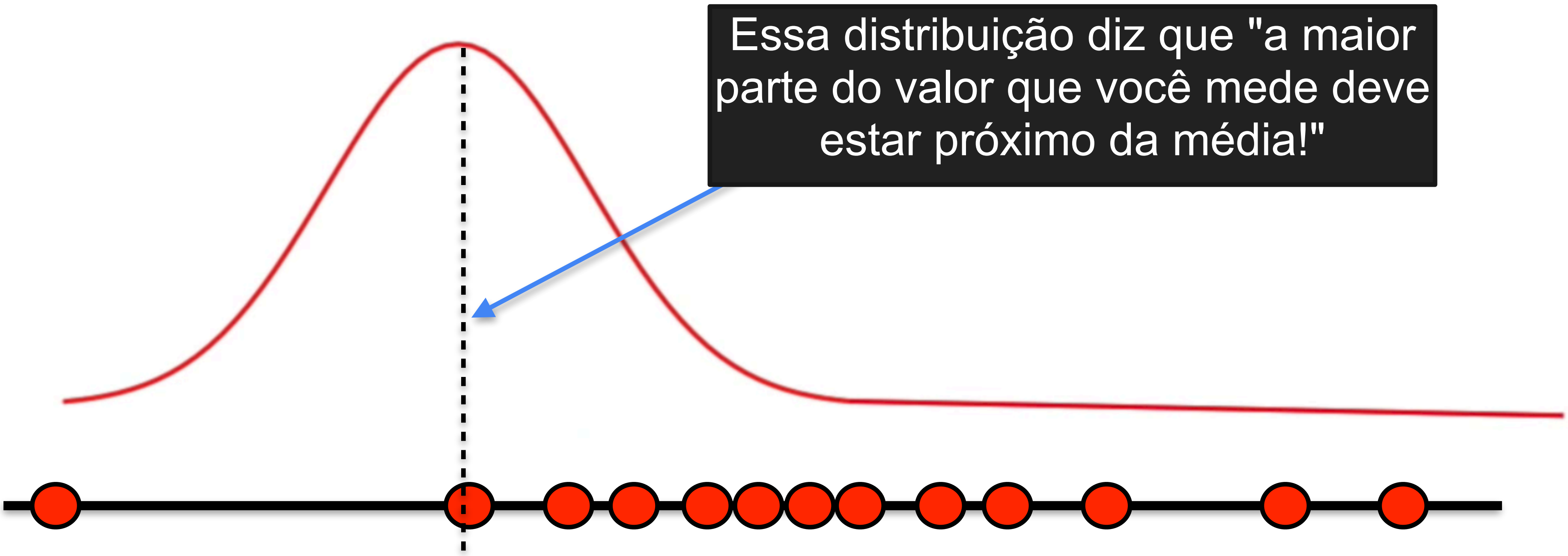
“Bem larga”

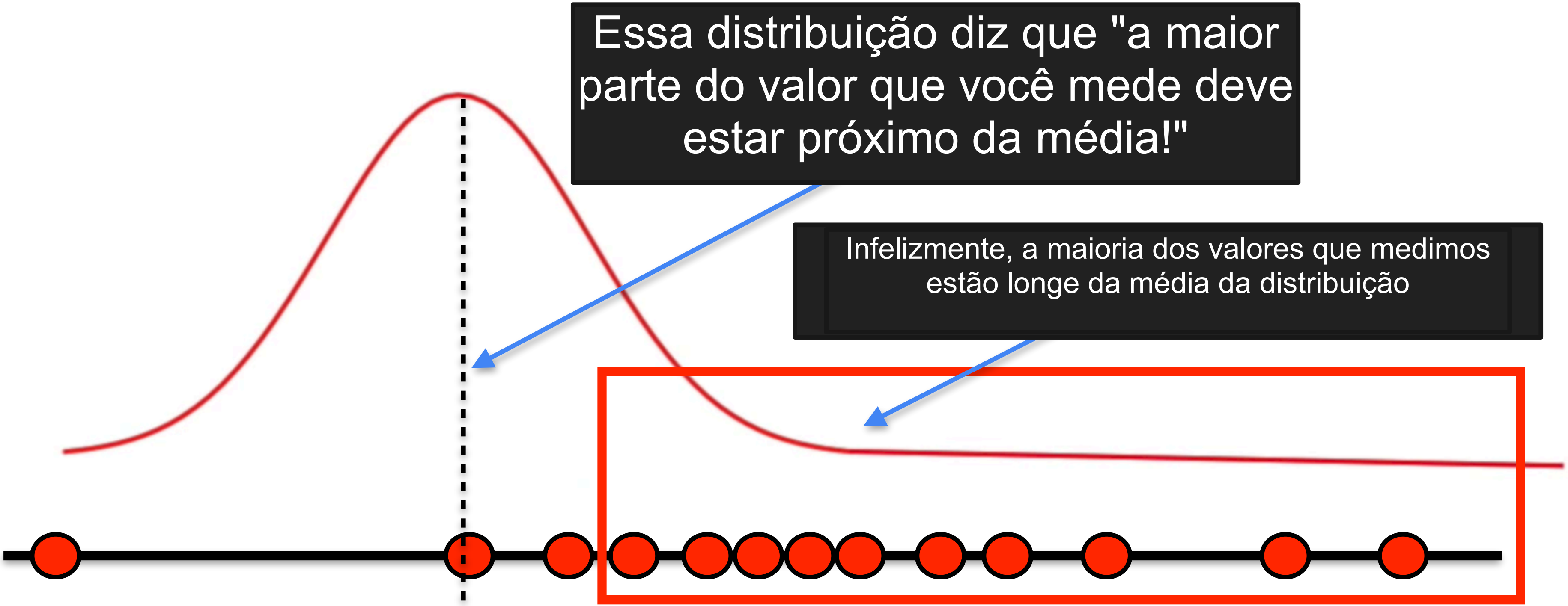
Uma vez que definimos a forma, temos que descobrir onde centralizar a distribuição.
Um local é melhor que outro

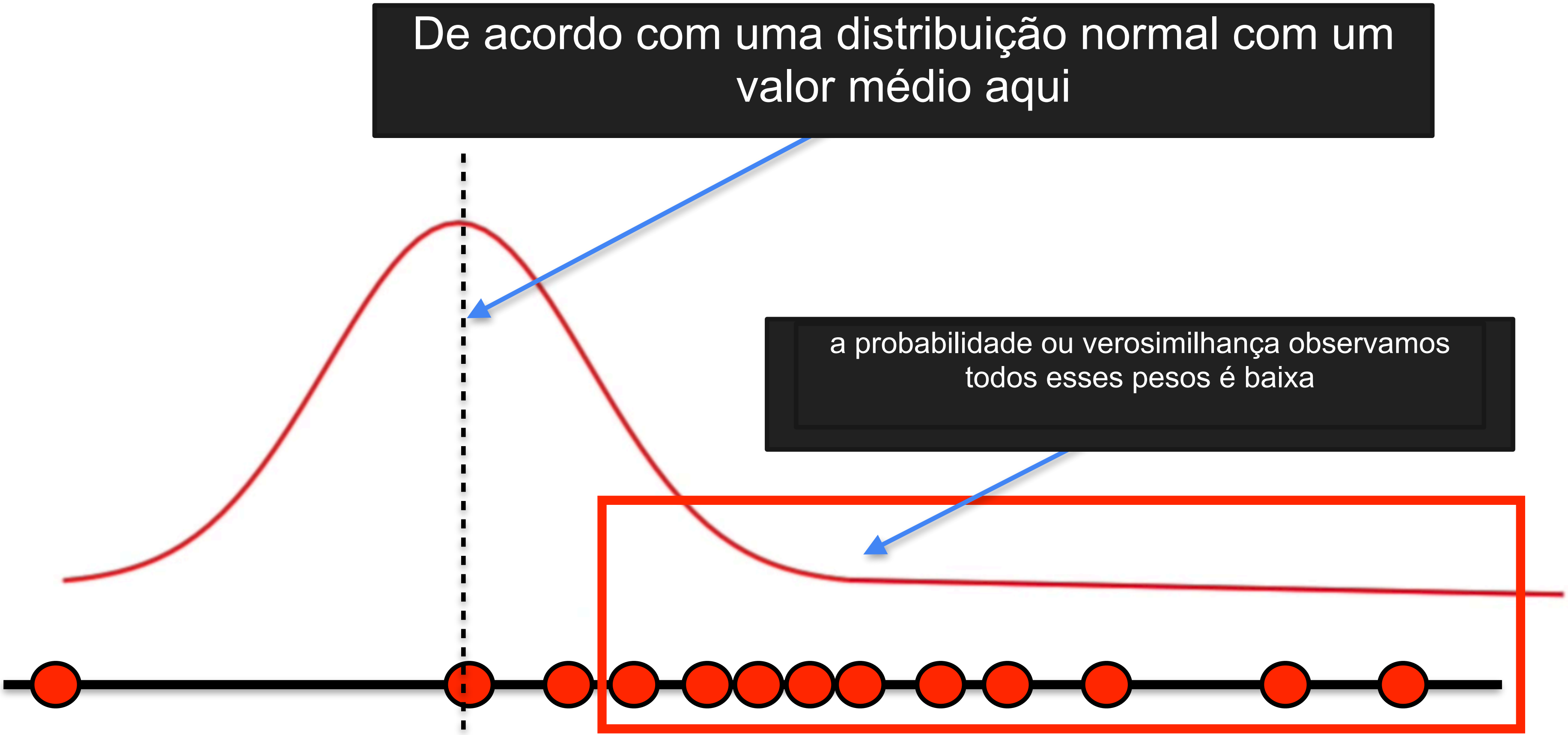


Antes de sermos muito técnicos, vamos escolher qualquer distribuição normal antiga e ver como ela se ajusta aos dados

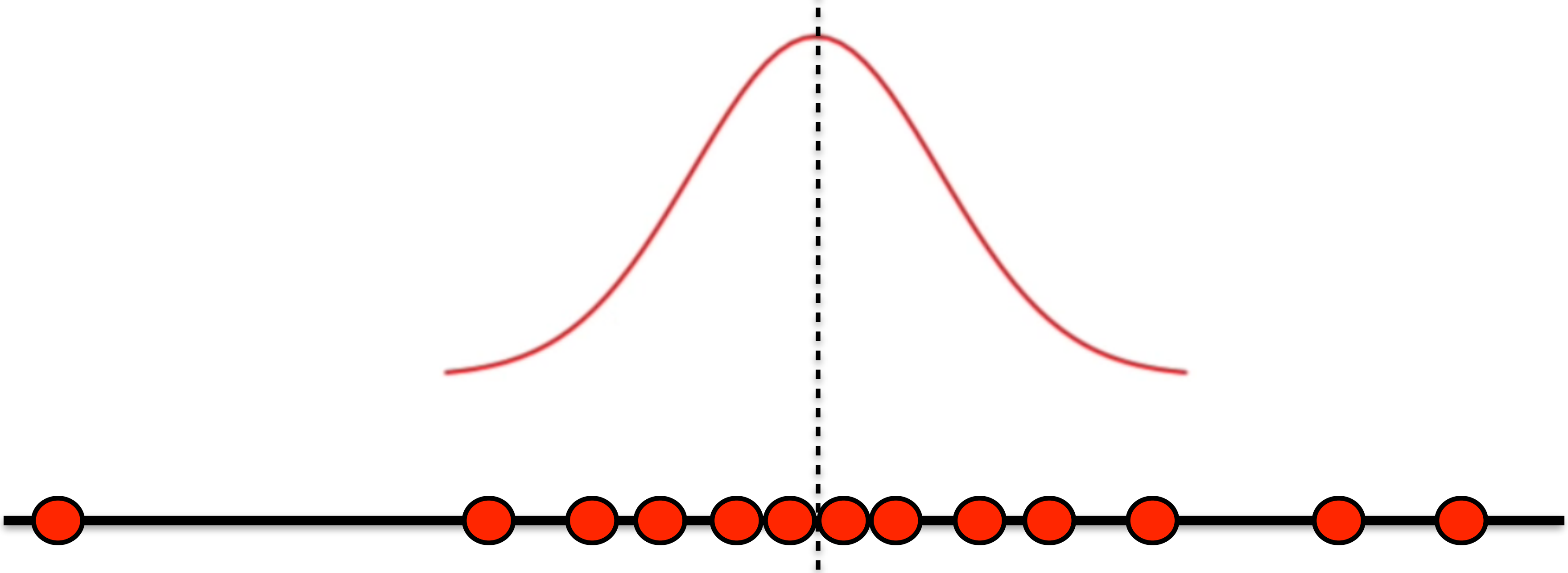



Essa distribuição diz que "a maior parte do valor que você mede deve estar próximo da média!"

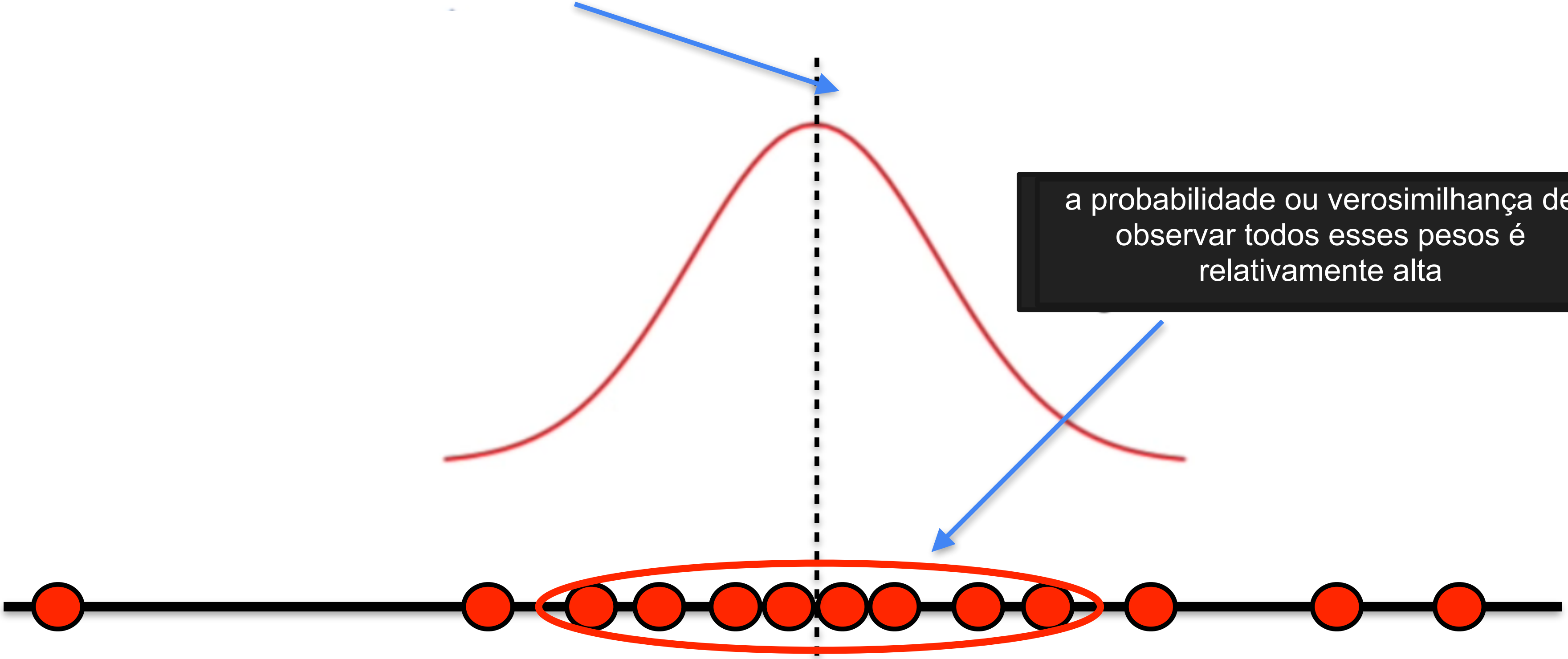

Essa distribuição diz que "a maior parte do valor que você mede deve estar próximo da média!"
Infelizmente, a maioria dos valores que medimos estão longe da média da distribuição


De acordo com uma distribuição normal com um valor médio aqui
a probabilidade ou verosimilhança observamos todos esses pesos é baixa

e se mudássemos a distribuição normal, de modo que sua média fosse igual ao peso médio?



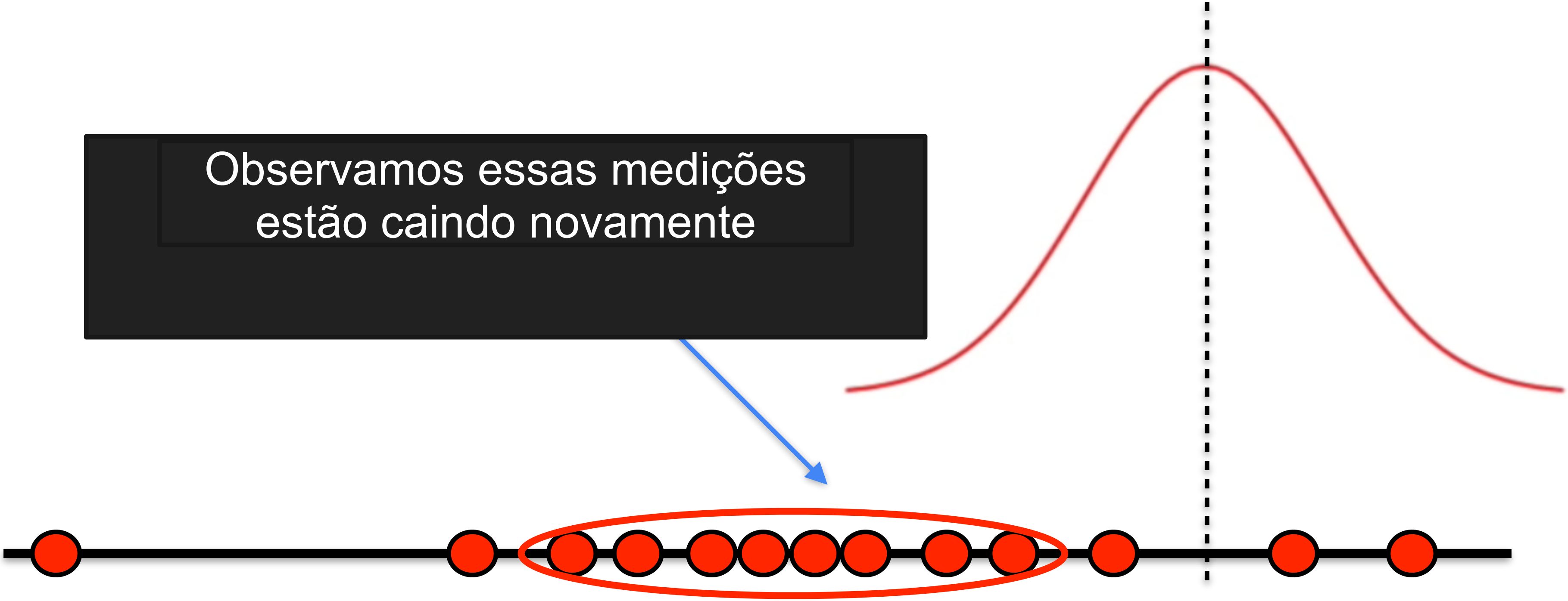
De acordo com uma distribuição normal com um valor médio aqui

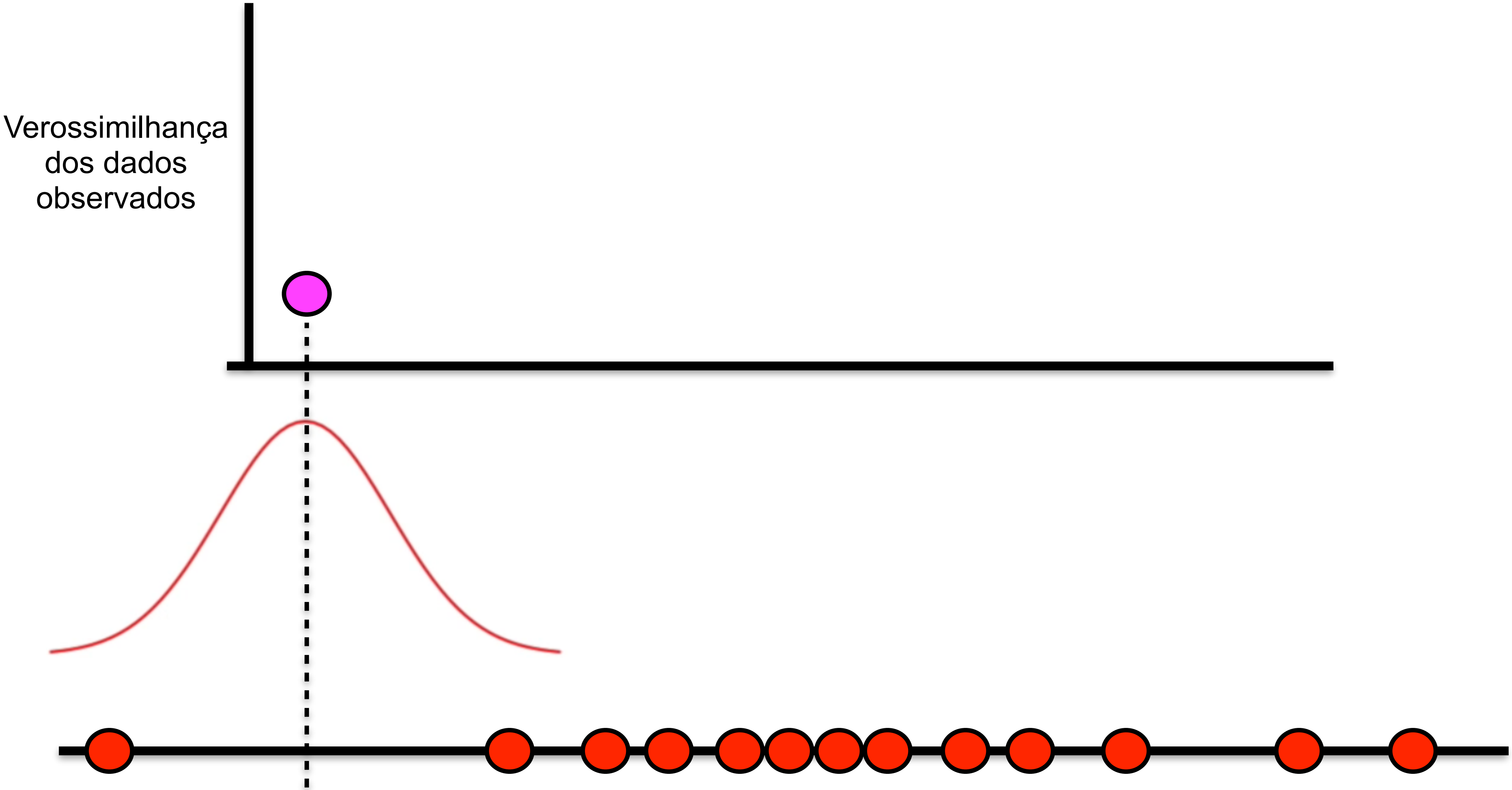


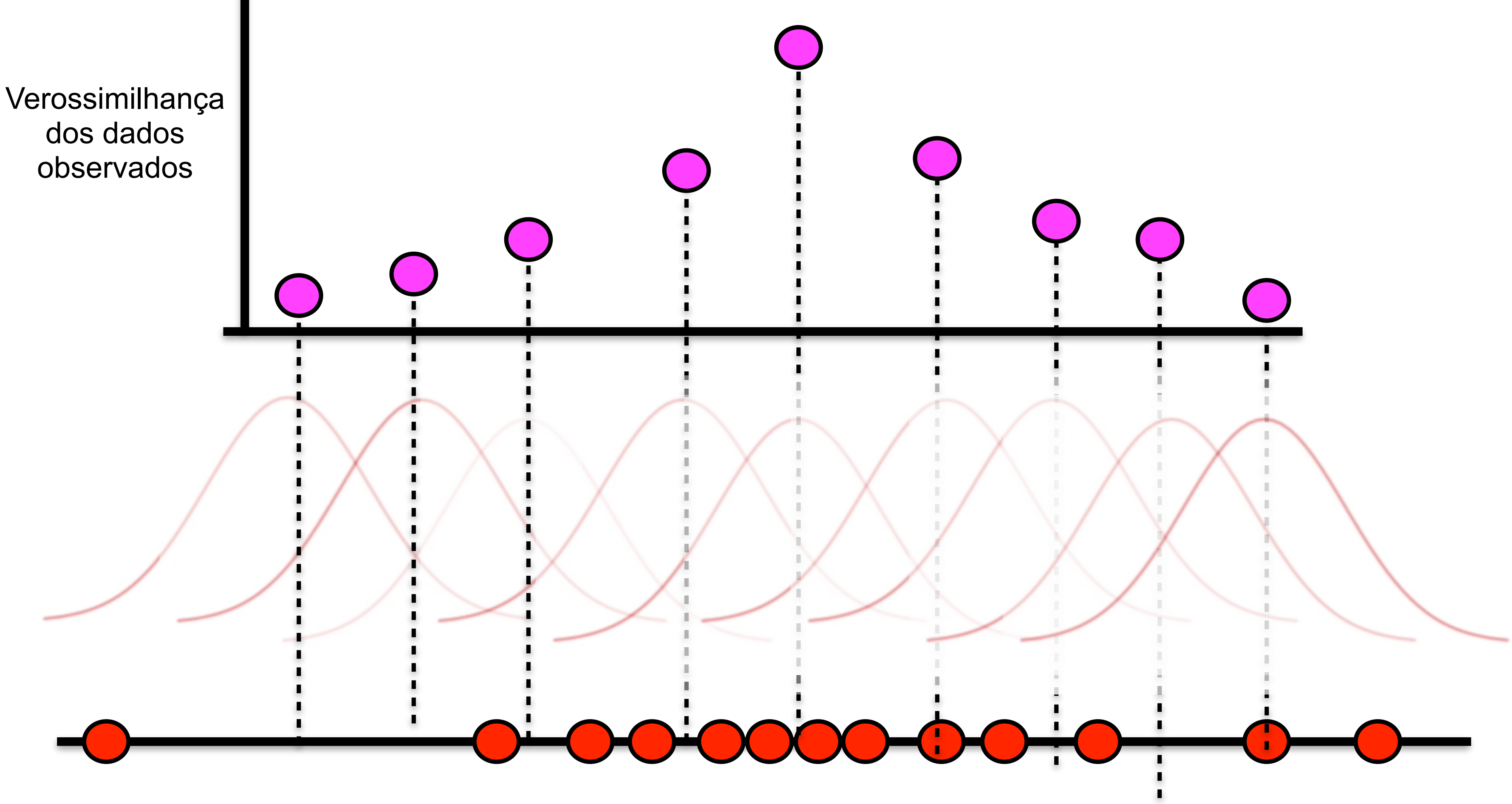
a probabilidade ou verosimilhança de observar todos esses pesos é relativamente alta

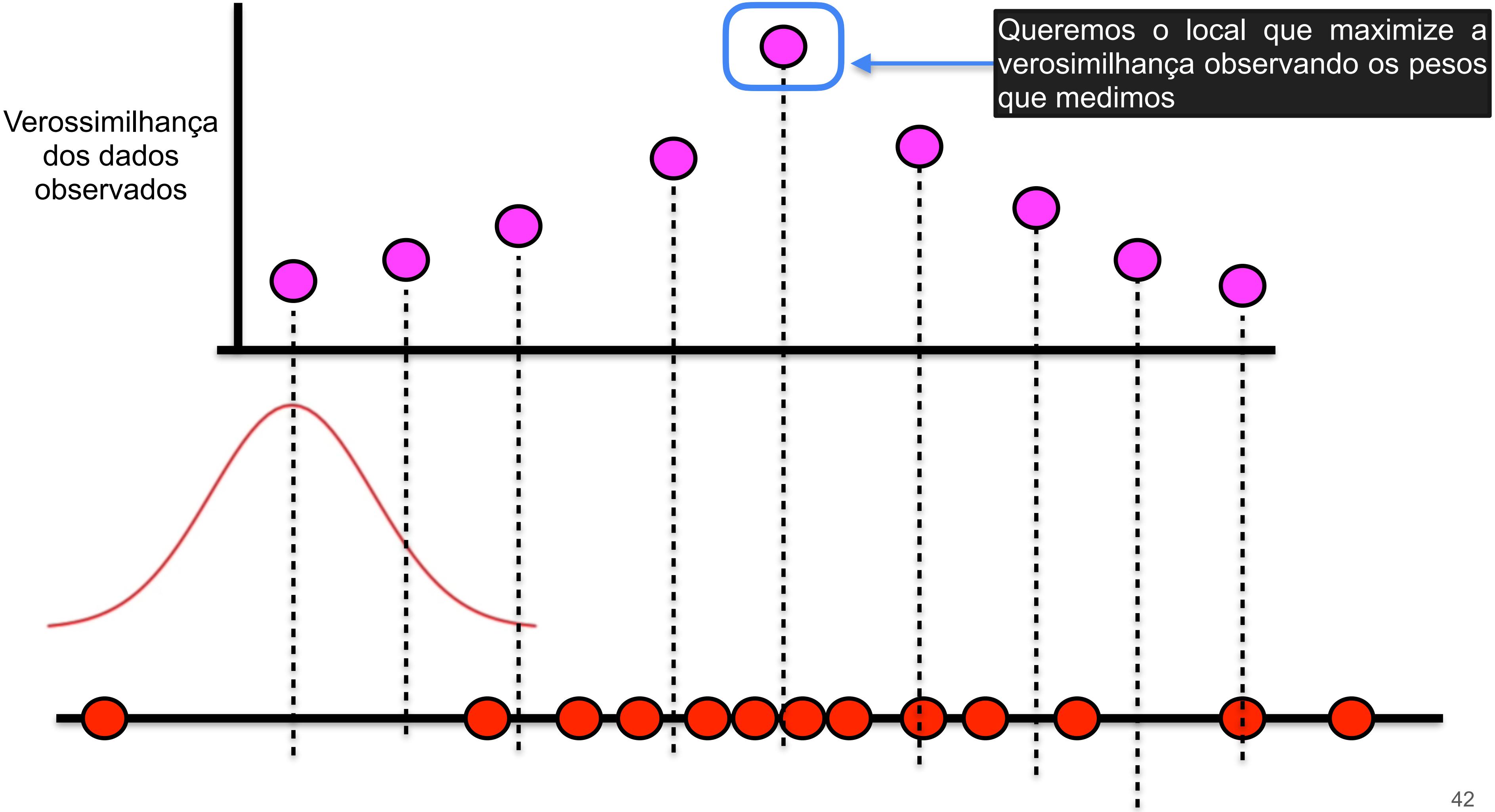
Se mantivéssemos mudando a distribuição normal

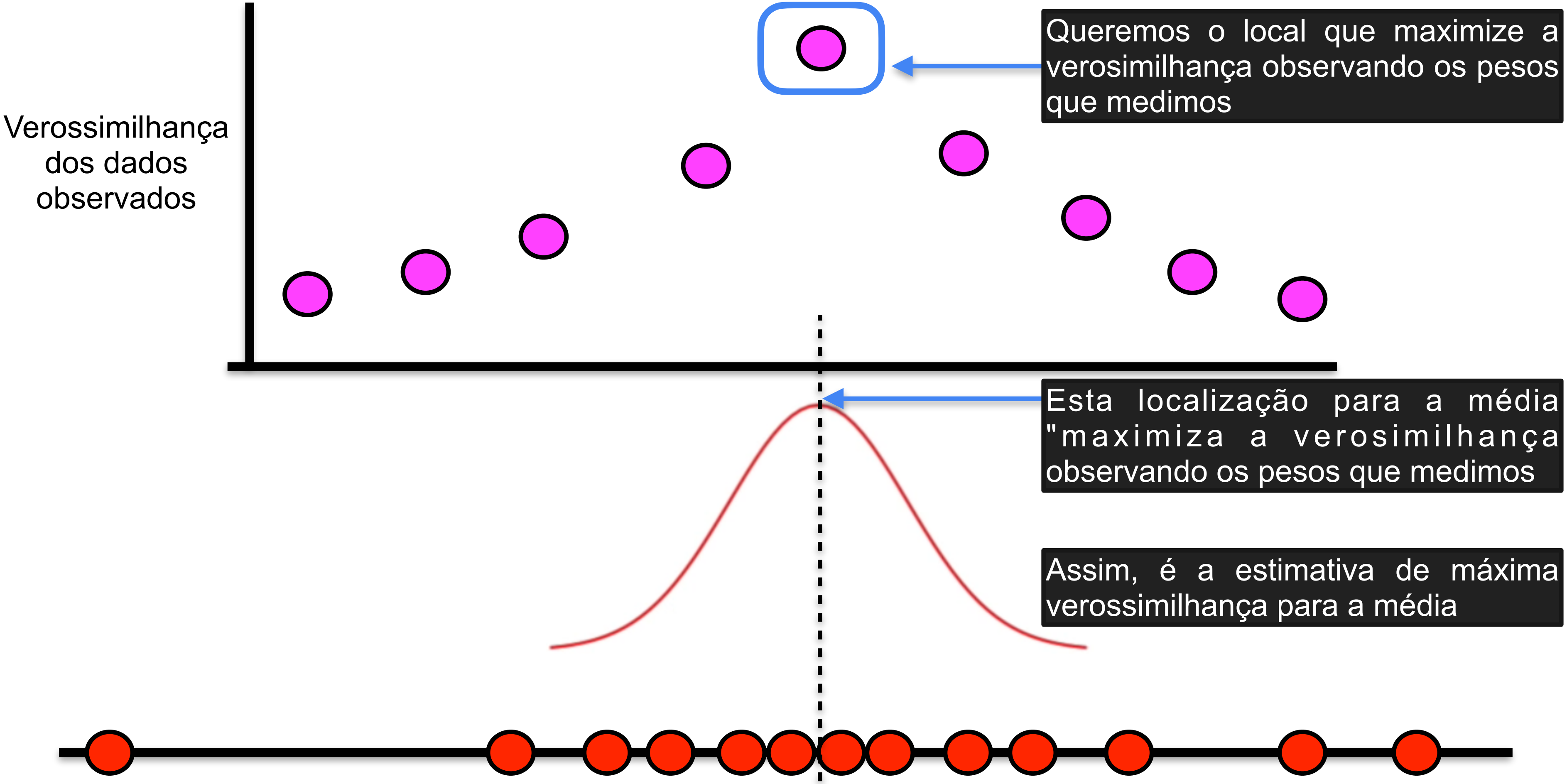
Observamos essas medições estão caindo novamente




Verossimilhança
dos dados
observados

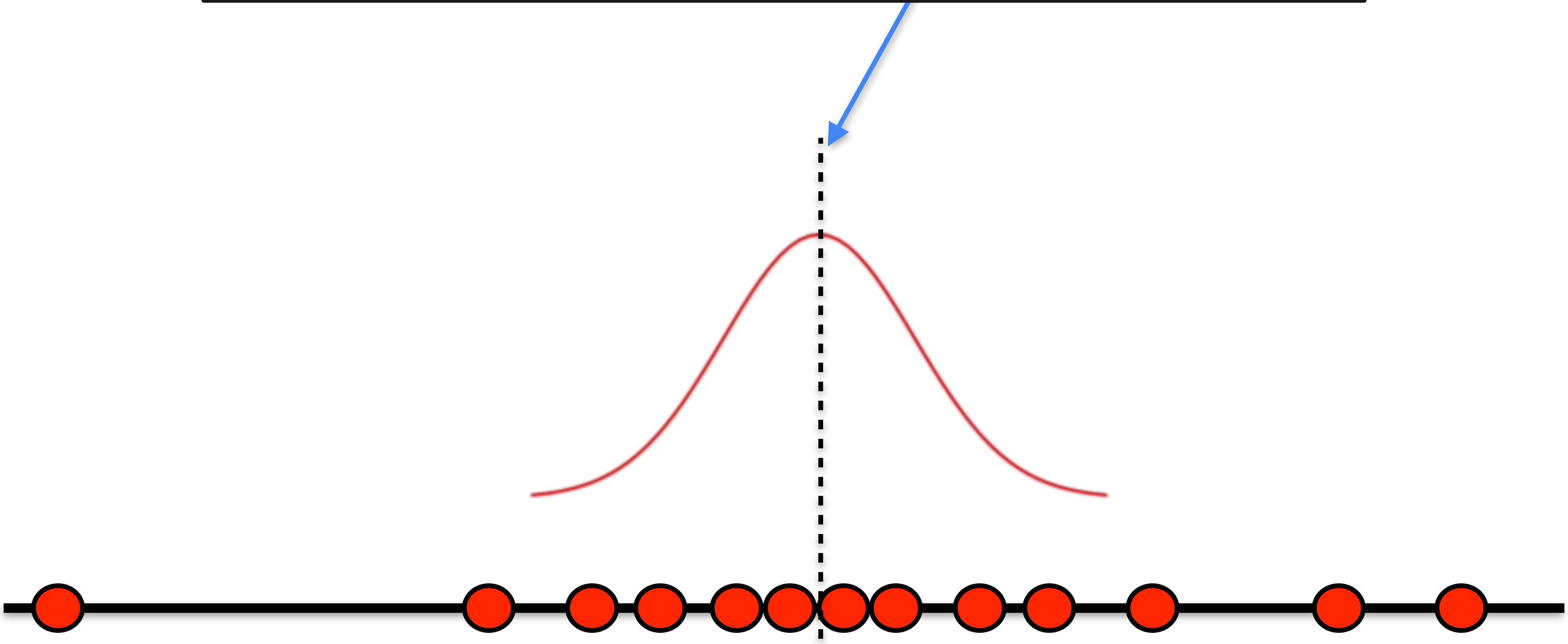

Verossimilhança
dos dados
observados


Verossimilhança
dos dados
observados
Queremos o local que maximize a verosimilhança observando os pesos que medimos

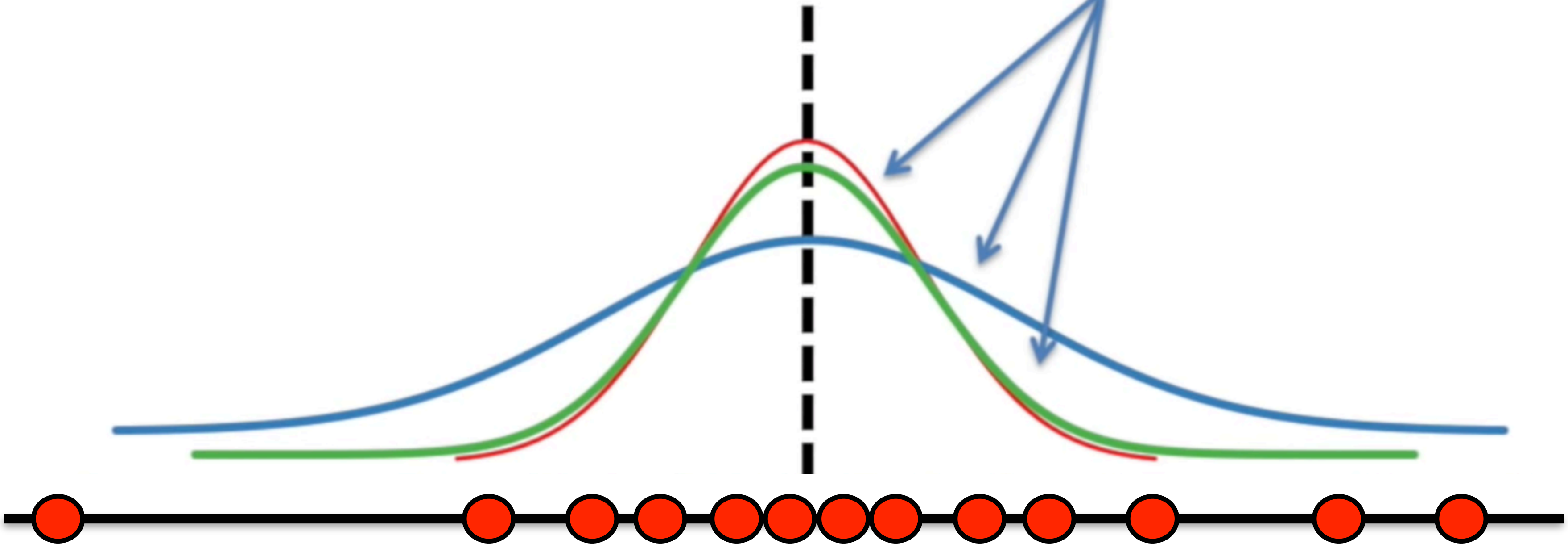


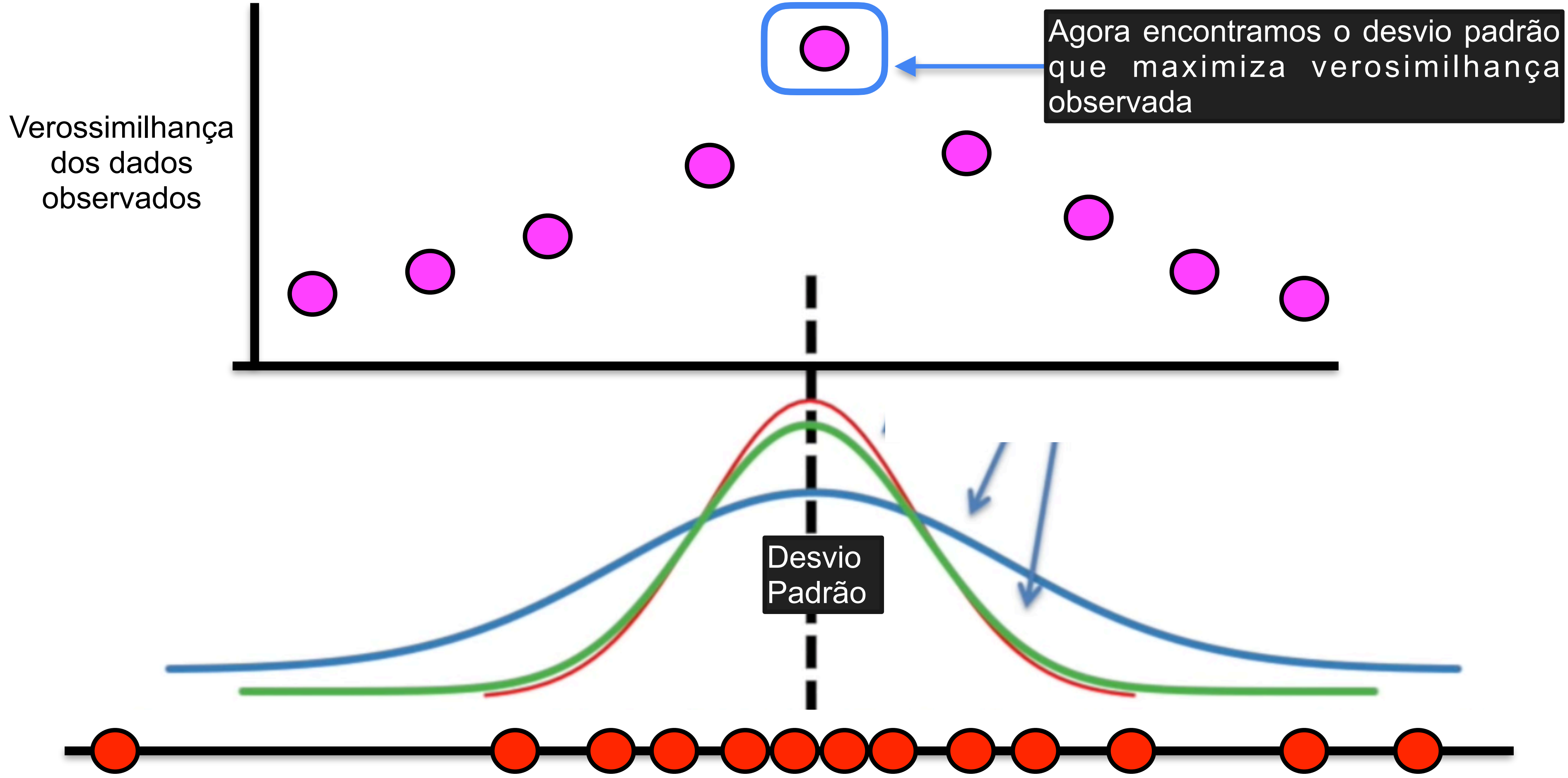
Queremos o local que maximize a verosimilhança observando os pesos que medimos
Verossimilhança dos dados observados
Esta localização para a média "maximiza a verosimilhança observando os pesos que medimos
Assim, é a estimativa de máxima verossimilhança para a média

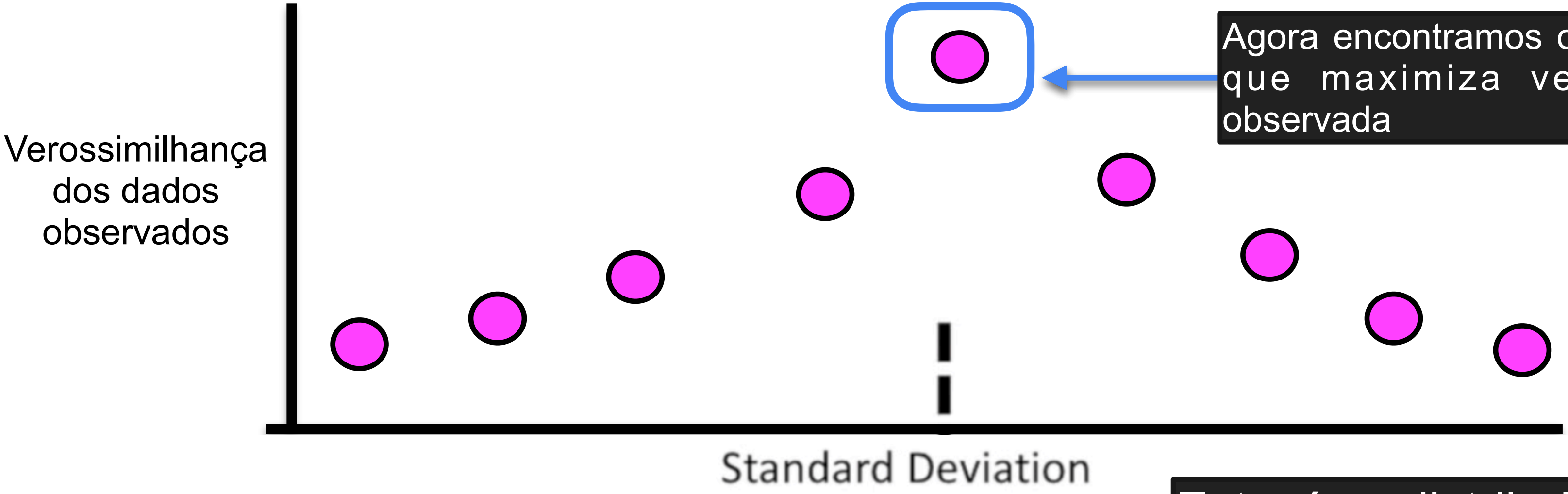
Excelente! Agora descobrimos a estimativa de máxima verossimilhança para a média!



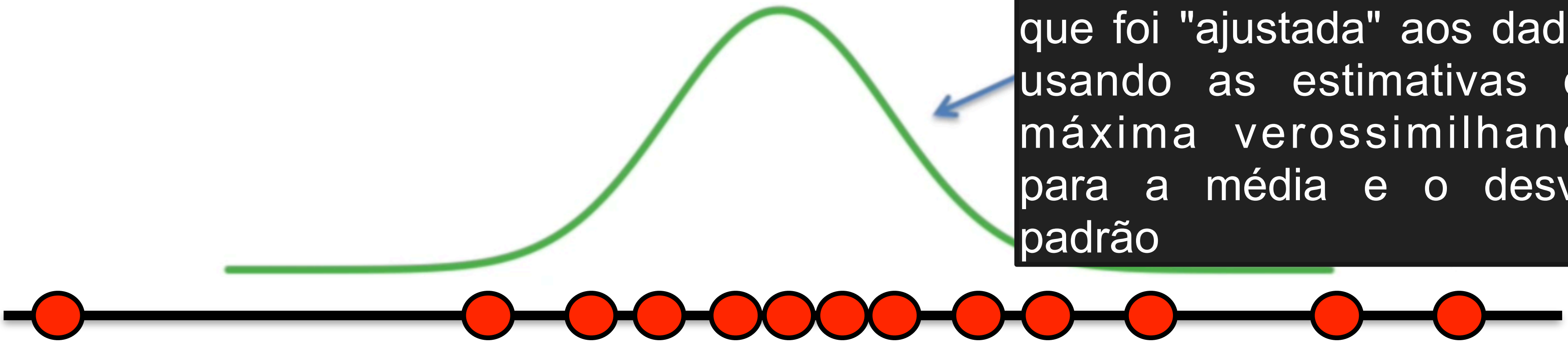
Agora temos que descobrir a "estimativa de máxima verossimilhança para o desvio padrão."




Verossimilhança
dos dados
observados
Agora encontramos o desvio padrão que maximiza verosimilhança observada
Desvio
Padrão

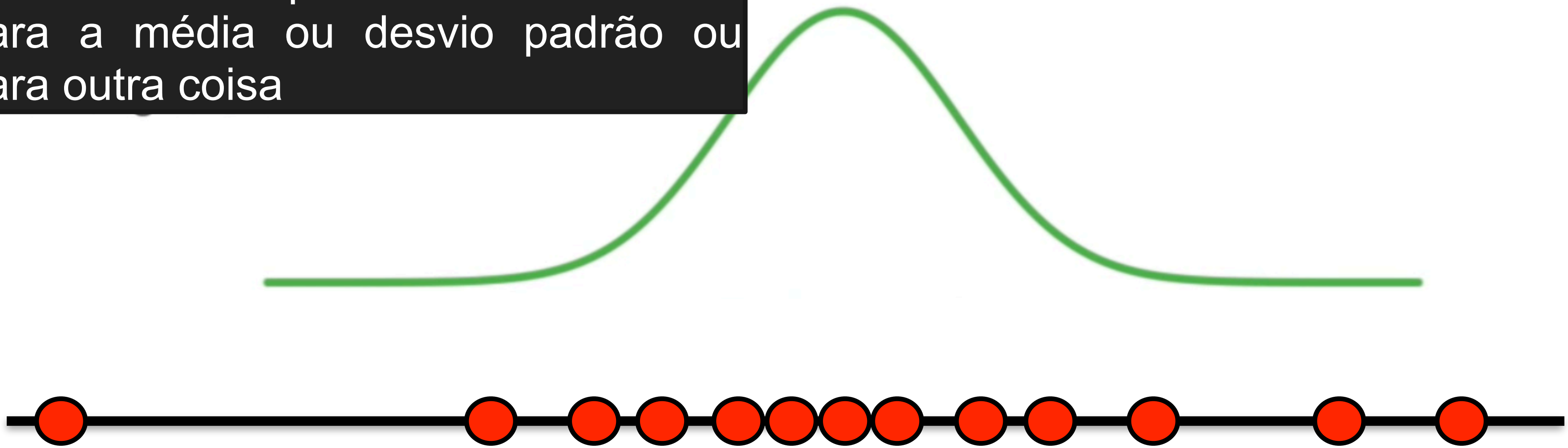


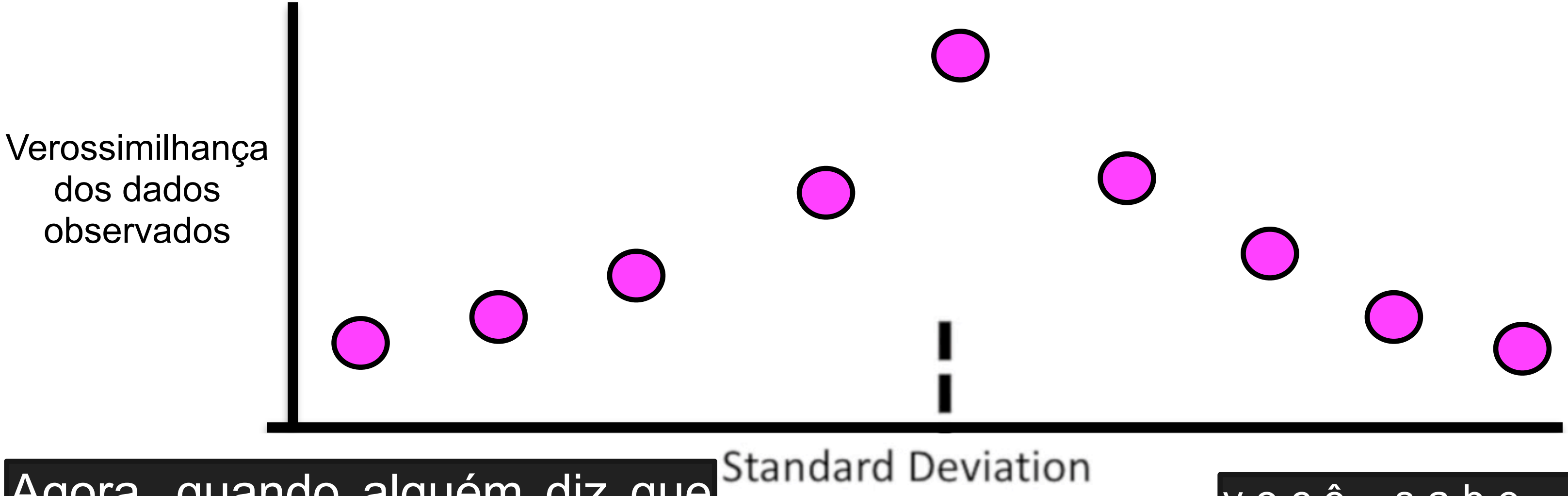
Agora encontramos o desvio padrão que maximiza verosimilhança observada



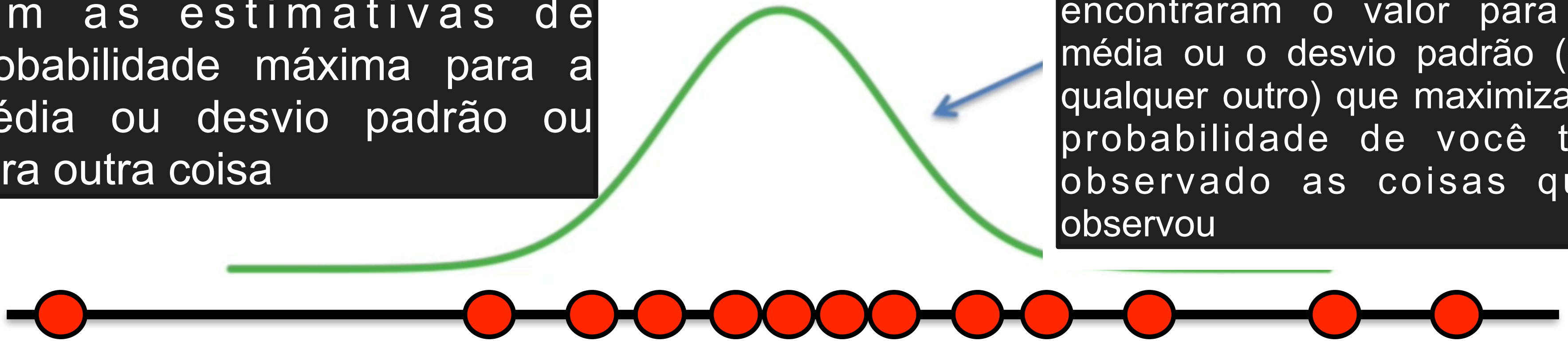
Esta é a distribuição normal que foi "ajustada" aos dados usando as estimativas de máxima verossimilhança para a média e o desvio padrão

Agora, quando alguém diz que tem as estimativas de probabilidade máxima para a média ou desvio padrão ou para outra coisa





Agora, quando alguém diz que tem as estimativas de probabilidade máxima para a média ou desvio padrão ou para outra coisa



você sabe que eles encontraram o valor para a média ou o desvio padrão (ou qualquer outro) que maximiza a probabilidade de você ter observado as coisas que observou

Sugestão de leitura e estudo complementar

- Seção 5.3, Páginas 98-119
- Os alunos que quiserem poderão ter um bônus se incluirem os exercícios no final do capítulo.

Testes Estatísticos

- Testes de significância e de hipóteses
- Testes unilaterais
    - p-value e níveis de significância
- Testes bilaterais
    - nível de confiança
    - Intervalo de confiança

Fonte: Notas do Prof. Vitor Oguri

Experimentos

dados $=\{x_1,x_2,...,x_N\} \Rightarrow \mu_{obs}$ (resultado)

Hipótese de nulidade

resultado esperado na ausência de qualquer outra causa que não seja o acaso.

(hipótese de nulidade) $H_0 : \mu = \mu_\circ$

(discrepância) $\mu_{obs \text{ diferente de }}$ $\mu_\circ$ (real ou casual?)

- Fisher: o objetivo de todo experimento é testar evidências contra a hipótese de nulidade de hipóteses

Fonte: Notas do Prof. Vitor Oguri

Testes estatísticos: significância e hipóteses

- pressupõem a escolha de uma distribuição de probabilidade (PDF), $\rho(\mu|\mu_\circ)$, para a distribuição dos dados ou para os valores de um parâmetro.
- testes de significância: consistência das observações (dados) com a hipótese de nulidade (não se considera hipóteses alternativas – Fisher)
- testes de hipótese: critérios para a rejeição ou aceitação da hipótese de nulidade (envolve hipóteses alternativas – Neyman-Pearson)

Fonte: Notas do Prof. Vitor Oguri

Testes unilaterais: p-value e nível de significância

- Proposta de Fisher
  - medida da evidência de que um valor observado de um parâmetro contrarie a hipótese de nulidade
  - probabilidade de que ocorra um resultado maior do que o valor observado

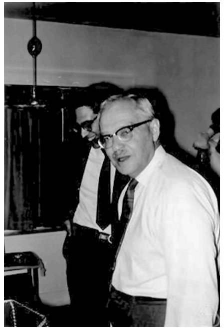
$$\begin{cases} p = P(\mu \geq \mu_{\text{obs}}) = \displaystyle\int_{\mu_{\text{obs}}}^{\infty} \rho(\mu|\mu_\circ)\, \mathrm{d}\mu \quad (p - \text{value}) \\ \\ p \to 0 \quad \big(\text{valor observado não compatível com hipótese de nulidade}\big) \end{cases}$$

jargão estatístico

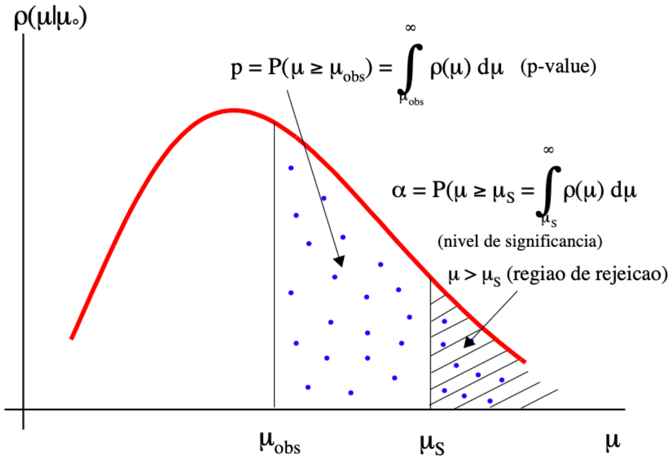
- $0{,}01 < p < 0{,}05$ – diz-se que a evidência contra a hipótese de nulidade não é significativa ao nível de 1%, mas é significativa ao nível de 5%

Fonte: Notas do Prof. Vitor Oguri

Testes unilaterais: p-value e nível de significância



- Proposta de Jerzy Neyman (1894-1981)
  - prefixa arbitrariamente um **nível de significância** (α) e, consequentemente, um valor limite para qualquer valor observado



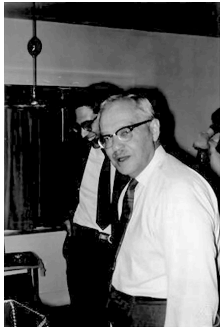
- $\alpha = P(\mu \geq \mu_S) = \int_{\mu_S}^{\infty} \rho(\mu|\mu_\circ)\, d\mu$
- $\mu \geq \mu_S$ (região de rejeição)

$$\mu_{obs} > \mu_S \iff p < \alpha$$

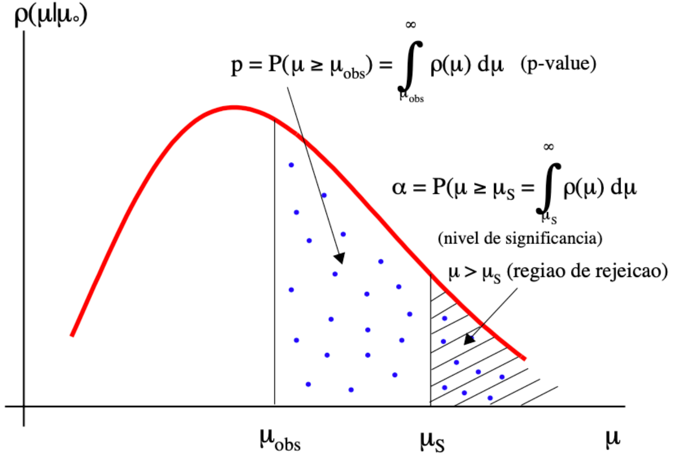
- hipótese de nulidade é rejeitada ao nível de $100 \times \alpha$ %
- discrepância estatisticamente significativa ao nível de $100 \times \alpha\%$

Fonte: Notas do Prof. Vitor Oguri

Testes unilaterais: p-value e nível de significância



- Proposta de Jerzy Neyman (1894-1981)
  - prefixa arbitrariamente um **nível de significância** (α) e, consequentemente, um valor limite para qualquer valor observado



$\alpha \to 1$ (grande) $\iff$ grande região de rejeição

$\Downarrow$

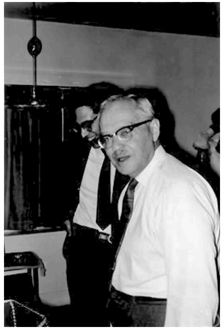
probabilidade alta de rejeição seja $H_\circ$ falsa ou verdadeira

rejeitar $H_\circ$ verdadeira $\Longrightarrow$ falsa descoberta (erro tipo I)

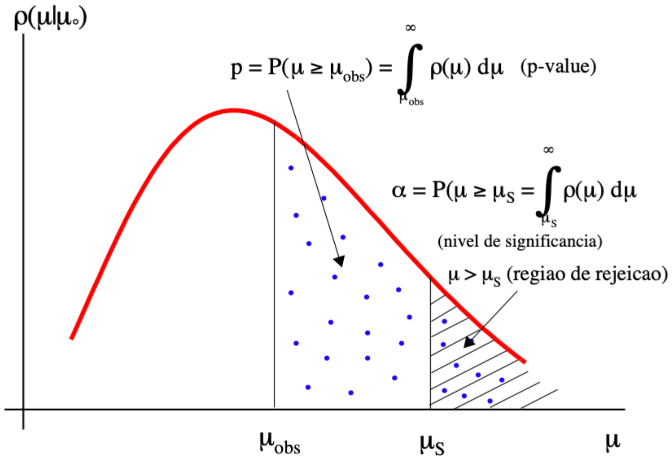
- a rejeição de uma hipótese de nulidade não significa de $H_\circ$ seja realmente falsa, significa apenas que a hipótese não é compatível com os dados
- outras hipóteses alternativas podem ser mais consistentes com os dados
- um único experimento não pode desacreditar uma "boa hipótese" (Fisher)

Fonte: Notas do Prof. Vitor Oguri

Testes unilaterais: p-value e nível de significância



- Proposta de Jerzy Neyman (1894-1981)
  - prefixa arbitrariamente um **nível de significância** (α) e, consequentemente, um valor limite para qualquer valor observado



$\alpha \to 0$ (pequeno) $\Longleftrightarrow$ grande região de aceitação

$\Downarrow$

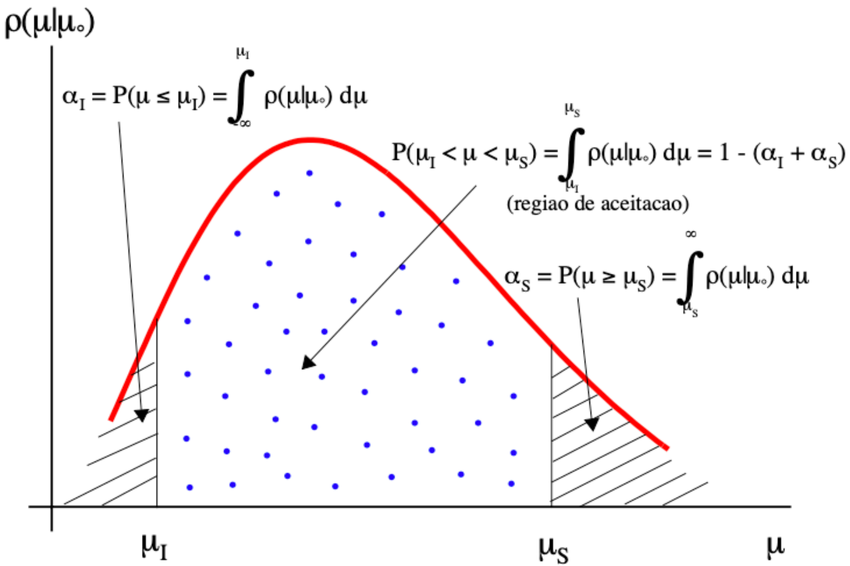
probabilidade alta de aceitação seja H_o falsa ou verdadeira

aceitar H_o falsa $\Longrightarrow$ refutar efeito real (erro tipo II)

- a aceitação de uma hipótese de nulidade significa apenas que H_0 e os dados são estatisticamente consistentes
- outras hipóteses alternativas podem ser igualmente
- consistentes não é um único resultado que estabelece uma "prova" estatística

Fonte: Notas do Prof. Vitor Oguri

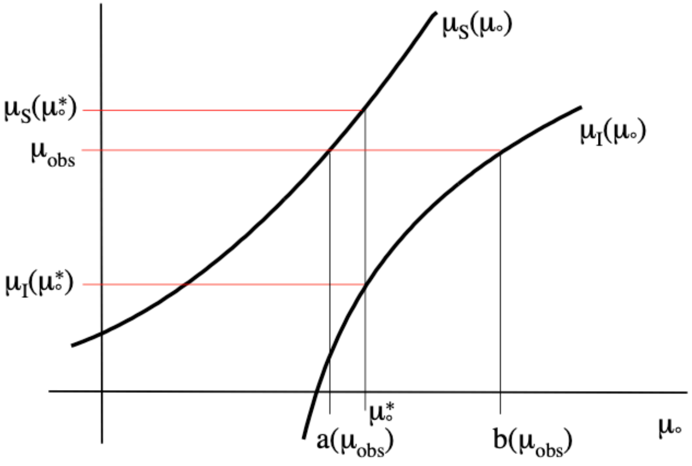
Testes bilaterais: Nível de confiança



- $(\mu_I, \mu_S) \rightarrow$ região de aceitação da hipótese de nulidade ao **nível de confiança** (CL) de $[1 - (\alpha_I + \alpha_S)] \times 100$ %

Fonte: Notas do Prof. Vitor Oguri

Testes bilaterais: Intervalo de confiança

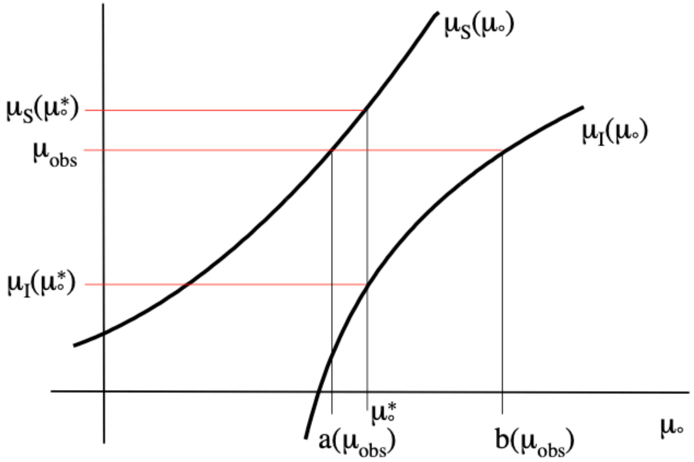


- para um dado valor esperado $\mu_\circ^*$ o valor observado μ_{obs} só estará em uma região de aceitação $\left(\mu_I(\mu_\circ^*), \mu_S(\mu_\circ^*)\right)$ definida por $\mu_\circ^*$ e pelo nível de confiança (CL) pré-fixado, se o intervalo $\left(a(\mu_{\text{obs}}), b(\mu_{\text{obs}})\right)$ determinado por μ_{obs} e pelo CL pré-fixado contiver o valor esperado $\mu_\circ^*$
- os intervalos (μ_I, μ_S) e (a, b) são probabilisticamente equivalentes, ou seja,
$\mu_I(\mu_\circ^*) < \mu_{\text{obs}} < \mu_S(\mu_\circ^*) \quad \Longleftrightarrow \quad a(\mu_{\text{obs}}) < \mu_\circ^* < b(\mu_{\text{obs}})$
- $P\left[\mu_I(\mu_\circ^*) < \mu_{\text{obs}} < \mu_S(\mu_\circ^*)\right] = P\left[a(\mu_{\text{obs}}) < \mu_\circ^* < b(\mu_{\text{obs}})\right]$
- (a, b) é um **intervalo de confiança** para $\mu_\circ$

- Um dado $\mu_\circ^*$ e um valor observado μ_{obs} só serão compatíveis se pertencerem, respectivamente, aos intervalos [(a, b) e (μ_I, μ_S)

Fonte: Notas do Prof. Vitor Oguri

Testes bilaterais: Intervalo de confiança - Interpretações Estatísticas



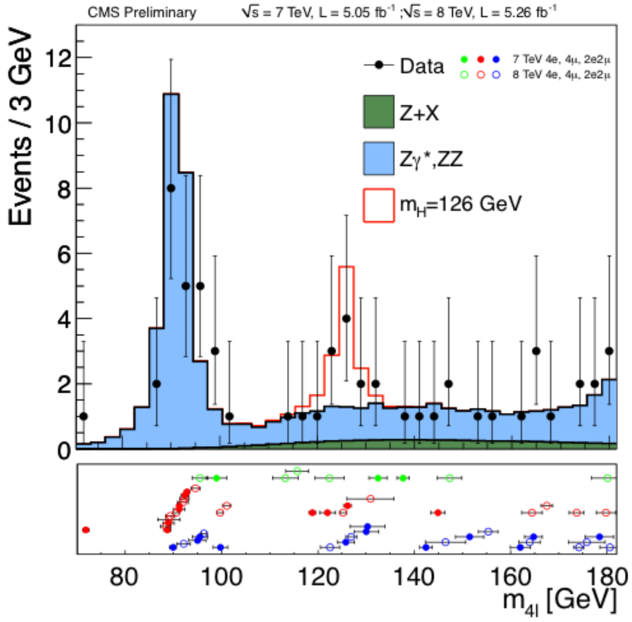
- Um dado $\mu_\circ^*$ e um valor observado μ_{obs} só serão compatíveis se pertencerem, respectivamente, aos intervalos $[(a, b)$ e (μ_I, μ_S)

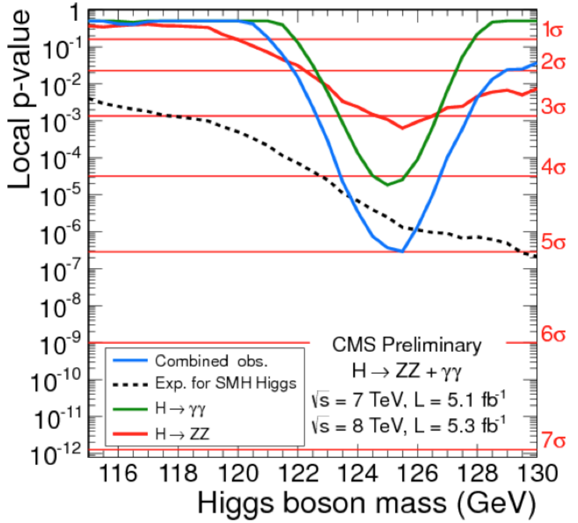
$$P(a<\mu_\circ <b)$$

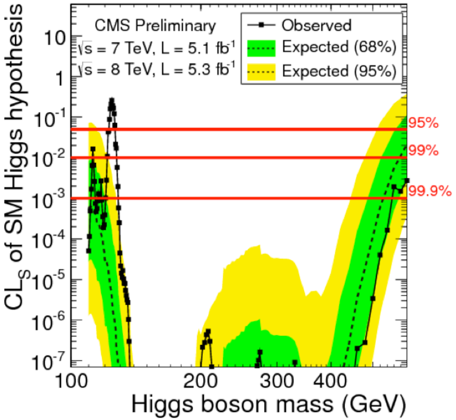
- fração de vezes que o intervalo (aleatório) (a, b) conterá o valor esperado do parâmetro $\mu_\circ$ (fixo – não-aleatório) (frequentista – Pearson-Neyman)
- probabilidade de que o valor esperado do parâmetro $\mu_\circ$ (aleatório) ocorra no intervalo (a, b) (subjetiva – bayesiana)

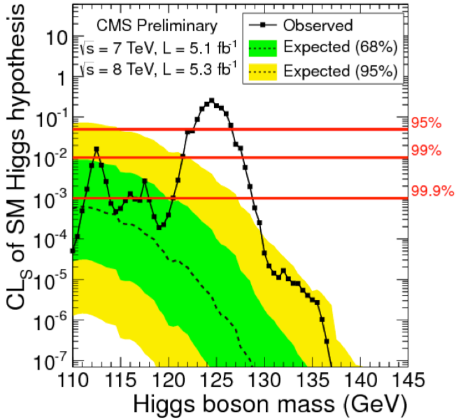
Fonte: Notas do Prof. Vitor Oguri

Significância e descoberta em FAE

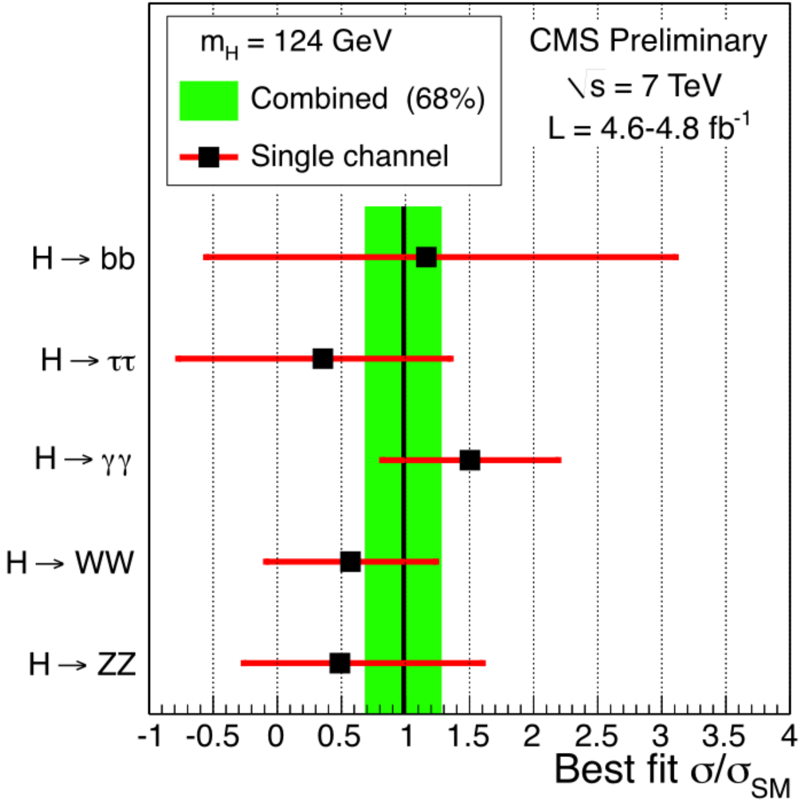


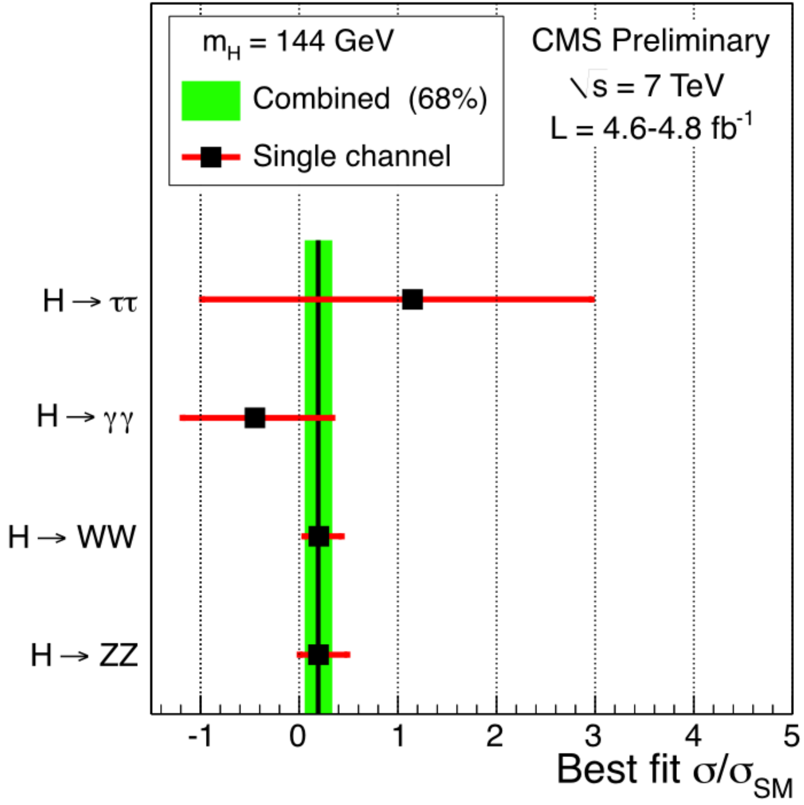




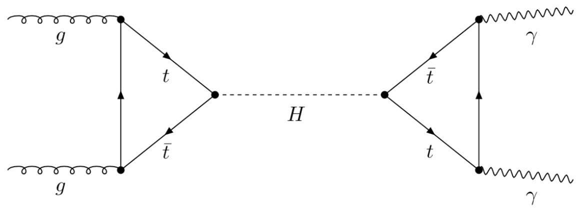


Significância e descoberta em FAE





Significância e descoberta em FAE



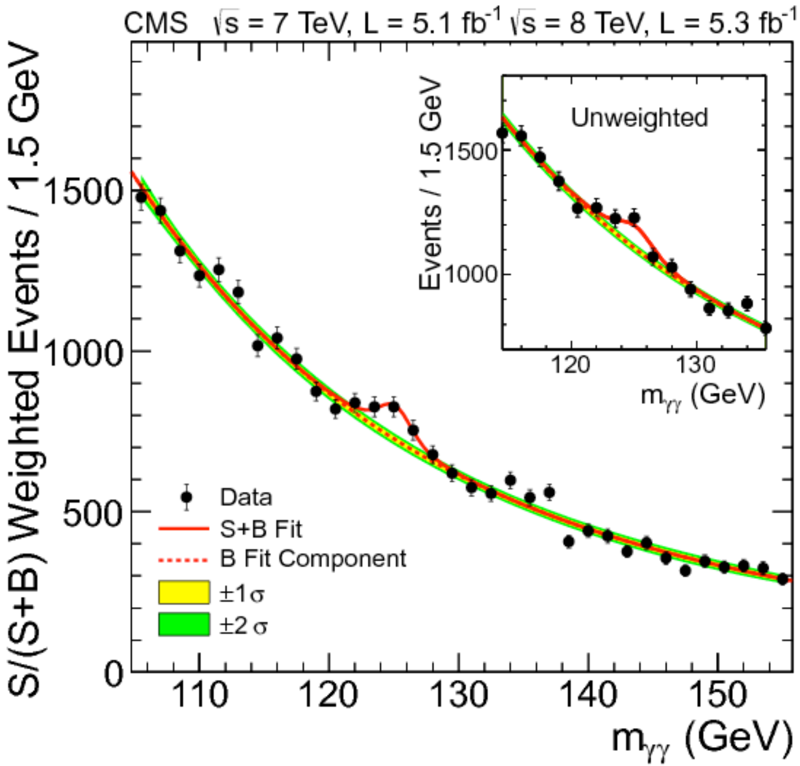


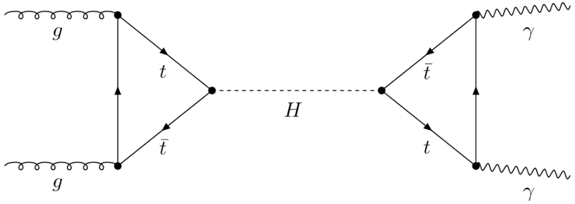
Tabela 6.3: Equivalência entre *p-value* e Z (sigma).

$Z(\sigma)$	p
1	0,15866
2	0,02275
3	0,00135
4	$3{,}1671\times10^{-5}$
5	$2{,}8665\times10^{-7}$
6	$9{,}8658\times10^{-10}$
7	$1{,}2798\times10^{-12}$
8	$6{,}1062\times10^{-16}$

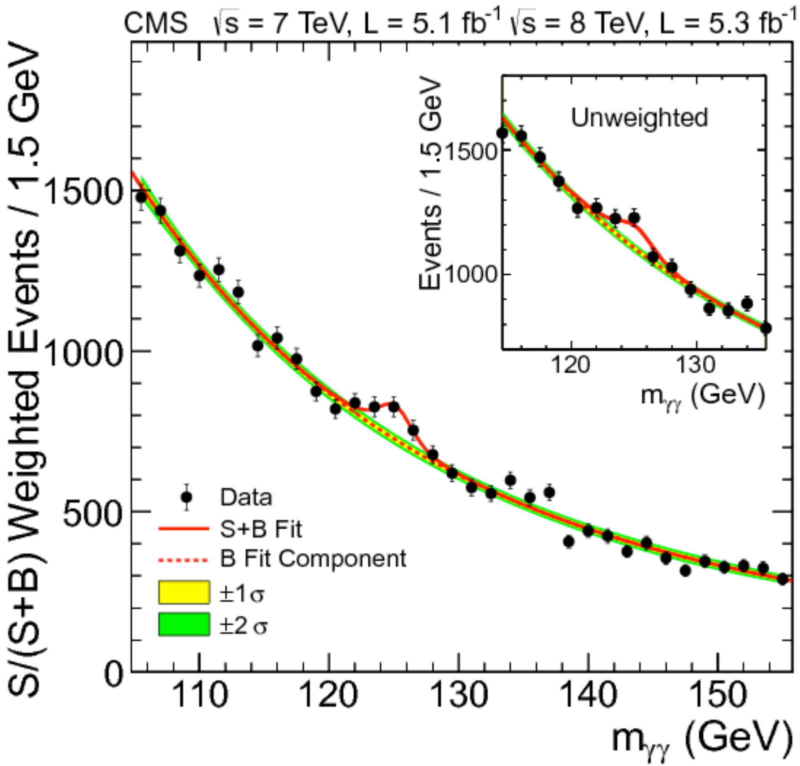
Seção 6.6, páginas 140-143 do livro do Vitor Oguri\
Sugestão de exercício bônus.

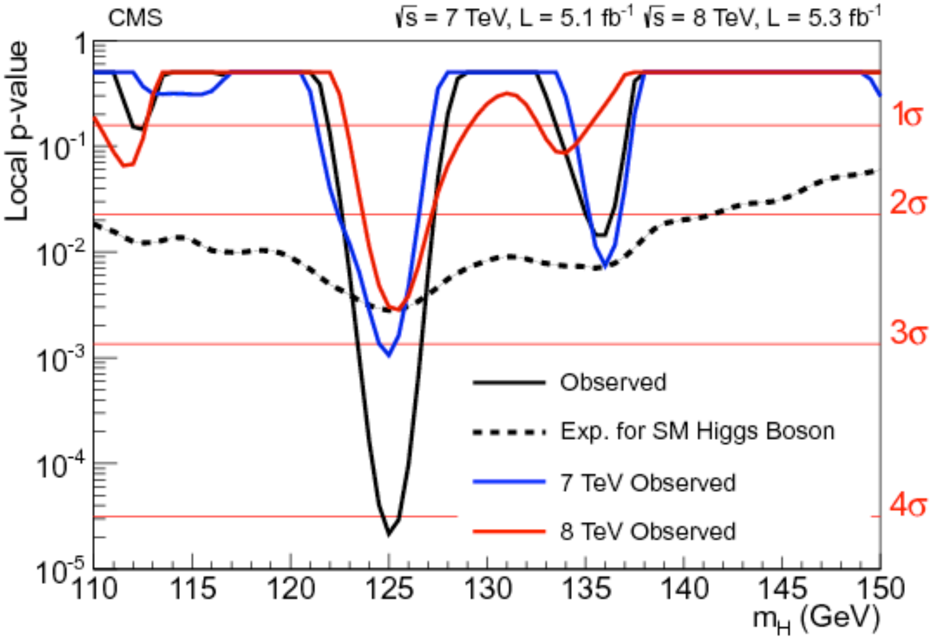
https://doi.org/10.1016/j.physletb.2012.08.021

Significância e descoberta em FAE



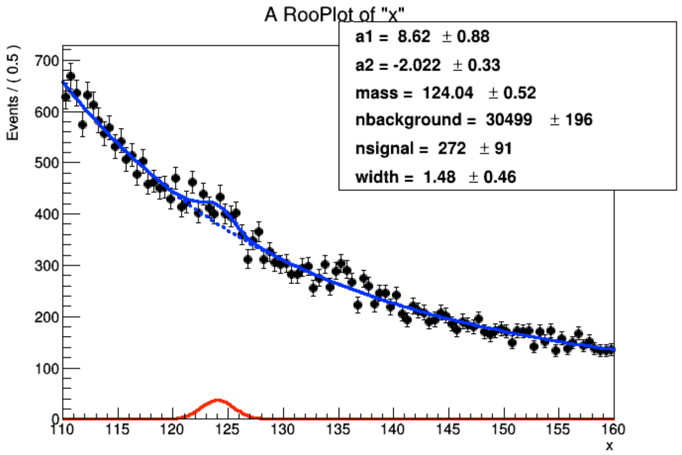
https://doi.org/10.1016/j.physletb.2012.08.021

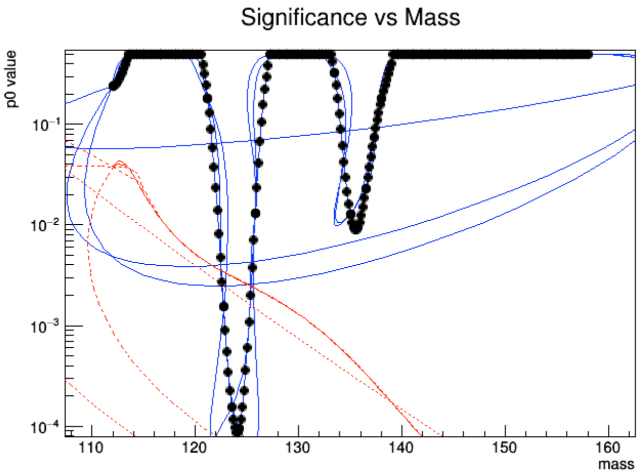




Notebooks

- Nas próximas aulas síncronas vamos apresentar exemplo de uma análise estatística para determinar os seguintes estudos para o bóson de Higgs decaindo em um par de fótons.
    - Variável de interesse
    - Limites de descoberta dos seus observáveis

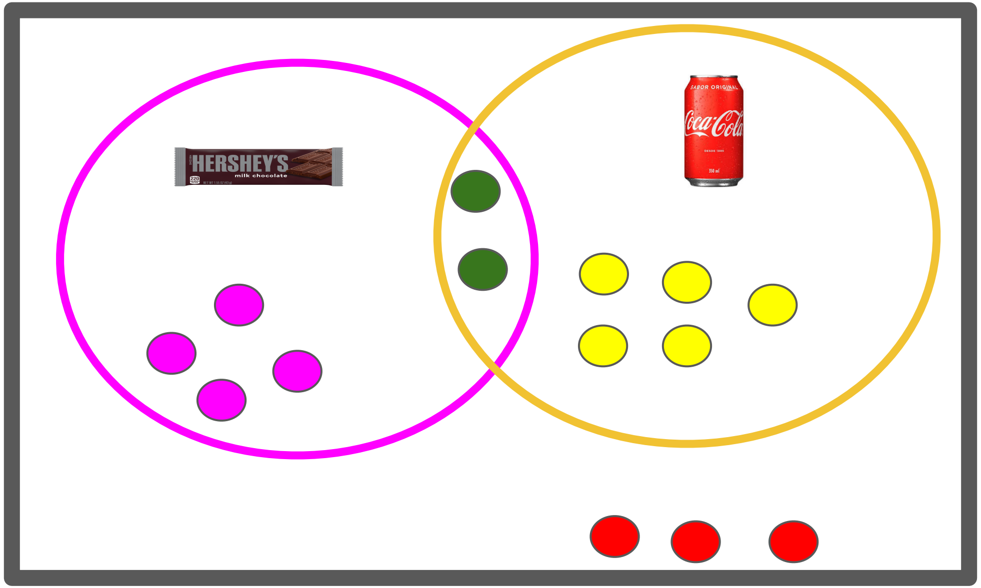


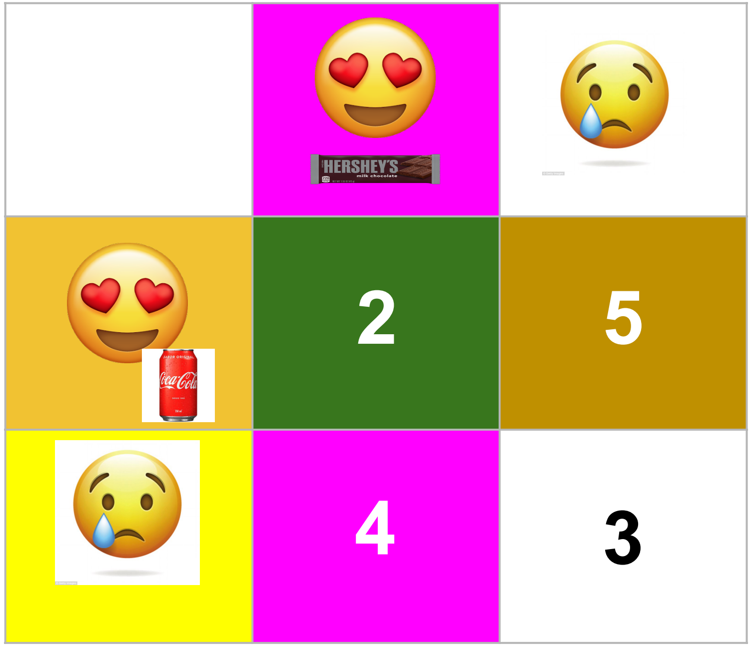


Backup slides

Probabilidade condicional

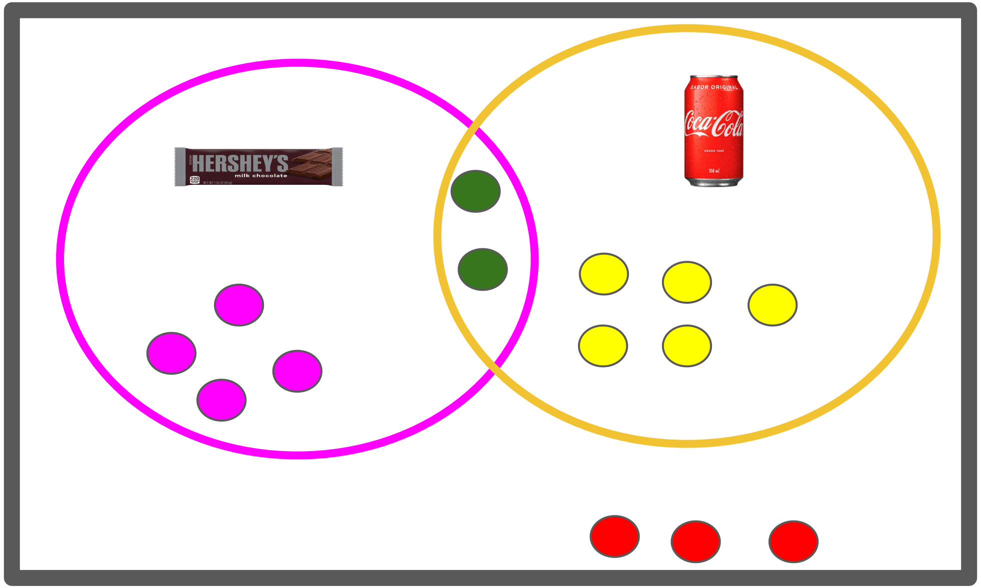
Probabilidade Condicional

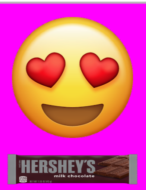




p(gosto de chocolate e refrigerante) = $\frac{2}{14}$

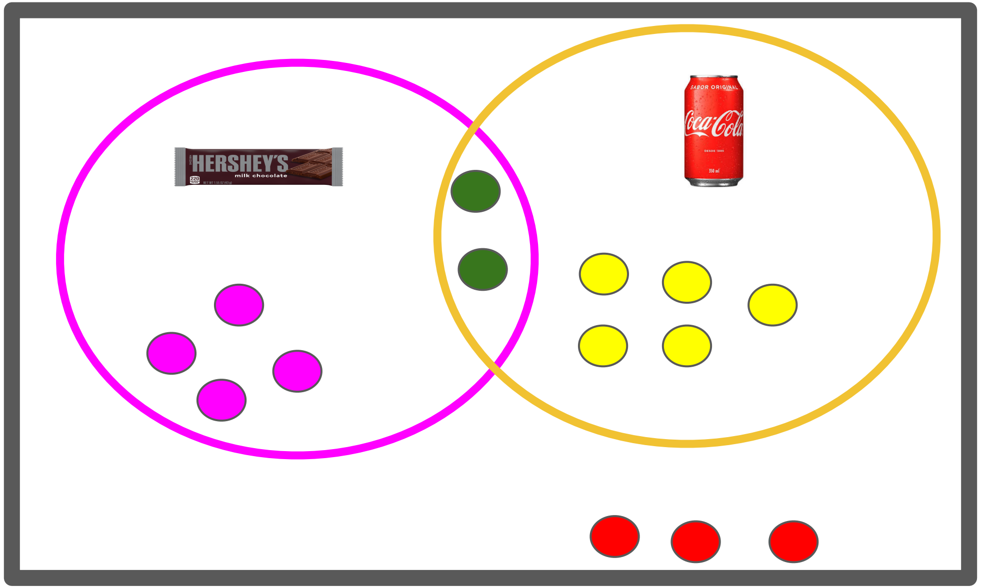
Probabilidade Condicional



		
	2 p = 2/14	5
	4	3

p(gosto de chocolate e refrigerante) = $\dfrac{2}{14} = 0,14$

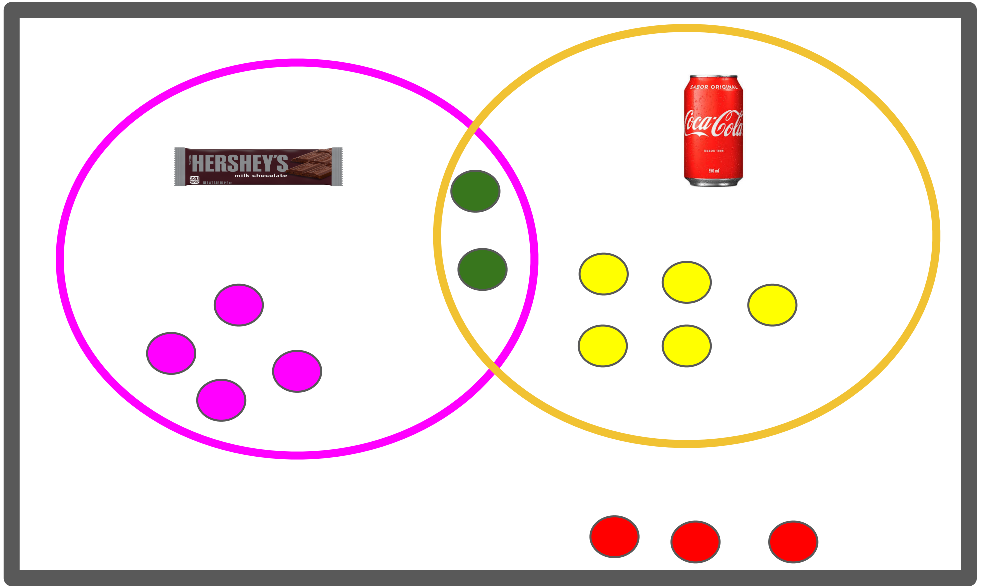
Probabilidade Condicional

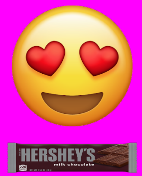
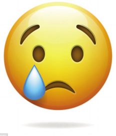
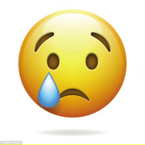


	😍 HERSHEY'S	😢
😍	2 p = 2/14	5
😢	4 p = 4/14	3

p(gosto de chocolate) = $\frac{4}{14} = 0,29$

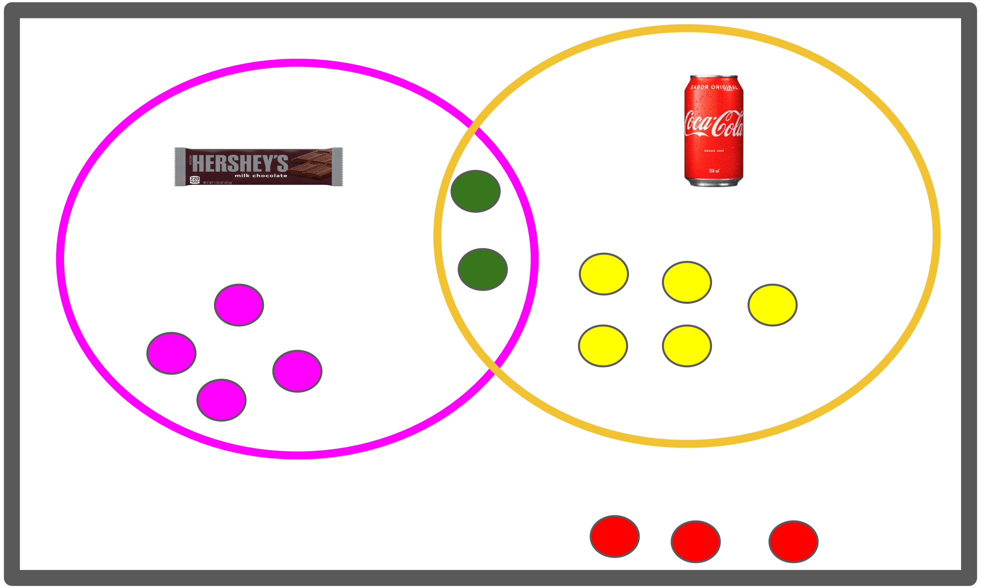
Probabilidade Condicional

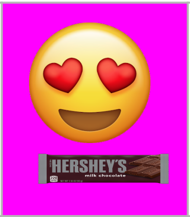
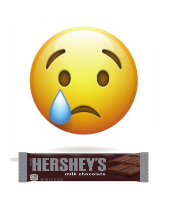
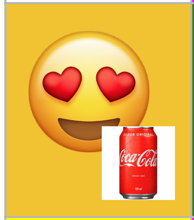


		
	2 p = 2/14	5 p = 5/14
	4 p = 4/14	3 p = 3/14

p(detesta chocolate e refrigerante) = $\frac{3}{14} = 0,21$

Probabilidade Condicional



			Total por linha
	2 p = 2/14	5 p = 5/14	2 + 5 p = 7/14
	4 p = 4/14	3 p = 3/14	4 + 3 p = 7/14
Total por coluna	2 + 4 p = 6/14		