

# **DELINEAMENTO E PLANEJAMENTO**

BIE 5793 - Princípios de Planejamento e Análise de Dados em Ecologia

# ETAPAS DO MÉTODO CIENTÍFICO

Scientific Method (1 serving)

1. Ask a question.
2. Formulate a hypothesis.
3. Perform experiment.
4. Collect data.
5. Draw conclusions.

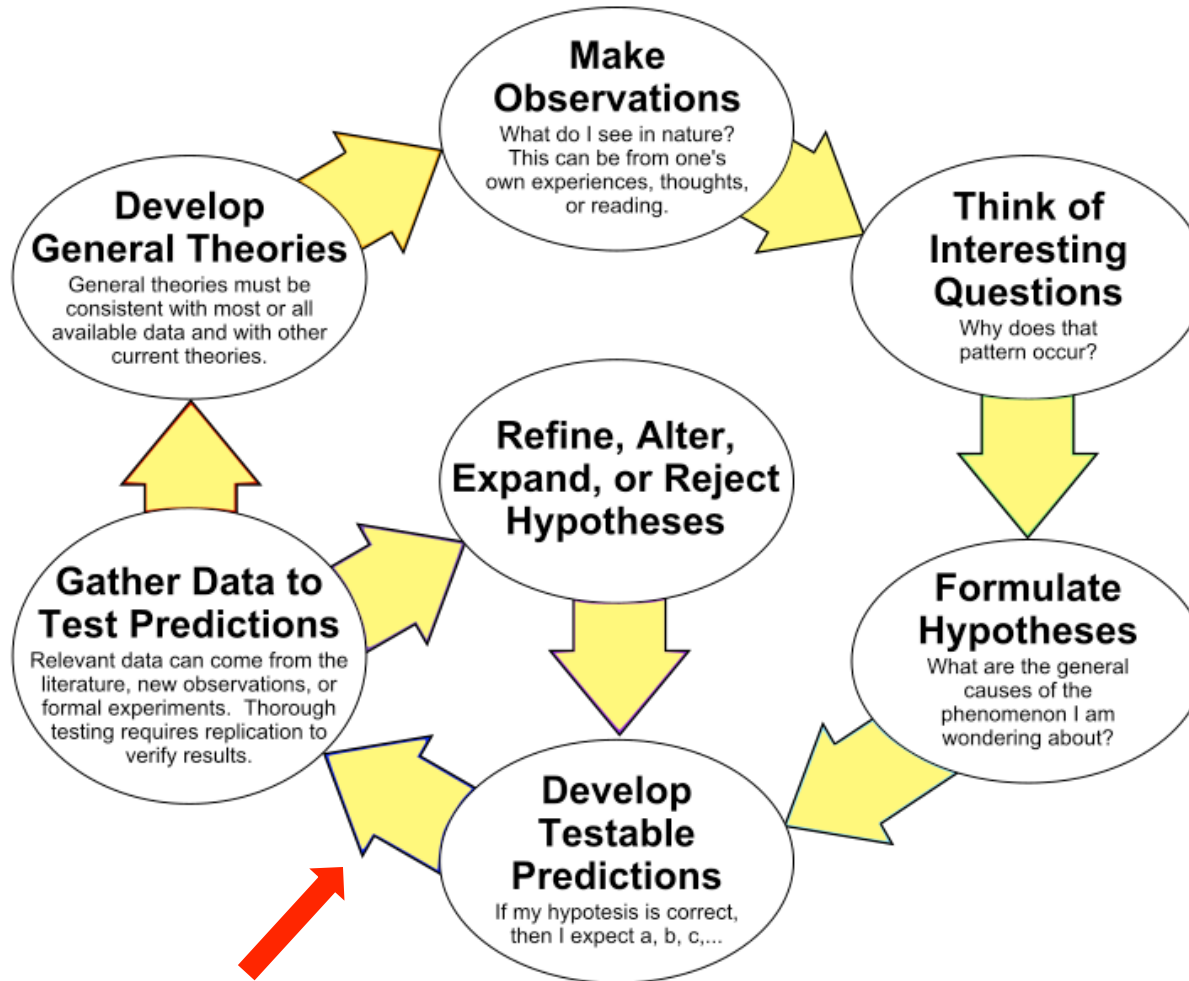
Bake until thoroughly cooked.

Garnish with additional observations.

*Too simple!*

# ETAPAS DO MÉTODO CIENTÍFICO

## The Scientific Method as an Ongoing Process



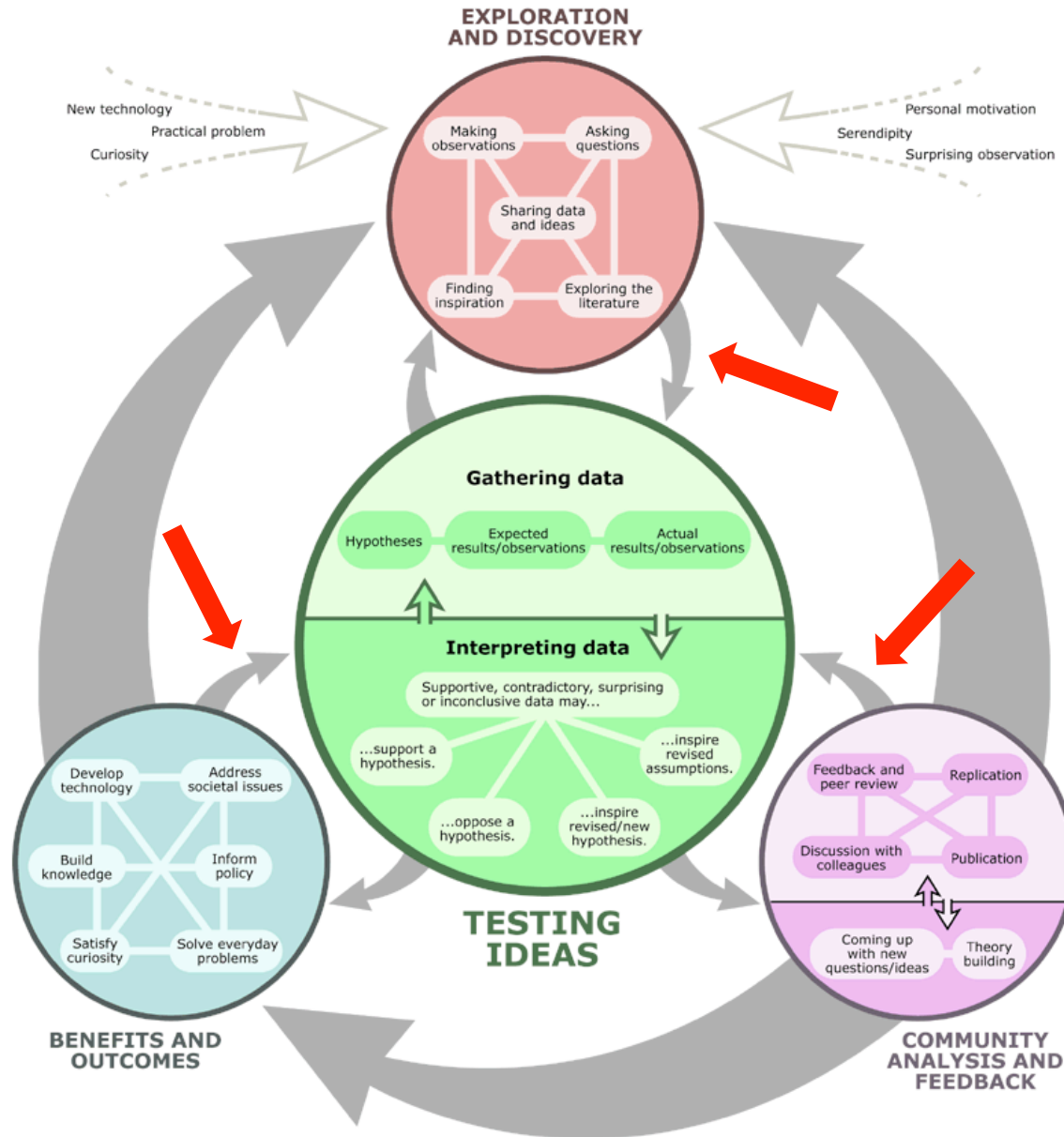
# ETAPAS DO MÉTODO CIENTÍFICO



**NÃO ABORDAM DELINEAMENTO E PLANEJAMENTO!!**

**VÃO DIRETO DAS HIPÓTESES/PREDIÇÕES PARA A COLETA DE DADOS**

# MESMO QUANDO O ESQUEMA NÃO É TRIVIAL



**QUANDO INCORPORA, PARECE UM PROCESSO SIMPLES E LINEAR**

**PREDIÇÕES**



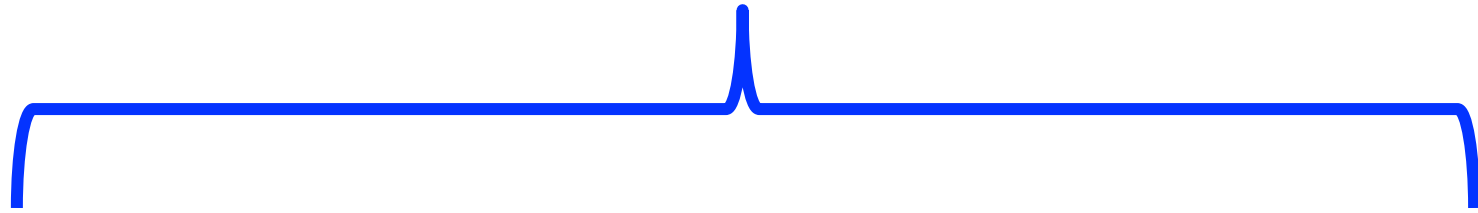
**DELINEAMENTO**

**COLETA DE DADOS**

**ANÁLISE**

**MAS, DELINEAMENTO É UMA ETAPA IMPORTANTE E COMPLEXA**

# Tudo depende da pergunta científica!



DELINEAMENTO



COLETA



ANÁLISE

Variáveis Operacionais  
(p. ex.: discreta x contínuas)

Tamanho de amostra  
Distribuição espacial  
Esforço amostral possível



Premissas

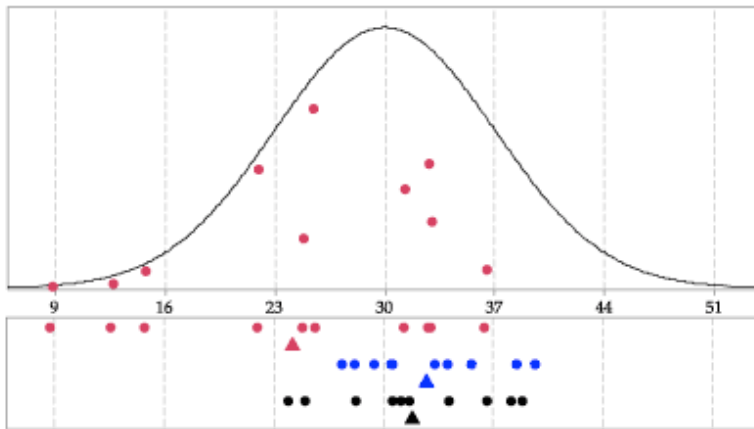
# POR QUE DELINEAMENTO É IMPORTANTE?

*"Designing an experiment properly will not only help you in analyzing data – it may determine **whether you can** analyze data at all!"*

Michael Palmer

## INCERTEZAS

INCERTEZAS INTRÍNSECAS  
DA  
POPULAÇÃO ESTATÍSTICA



**VARIÁVEIS ALEATÓRIAS**

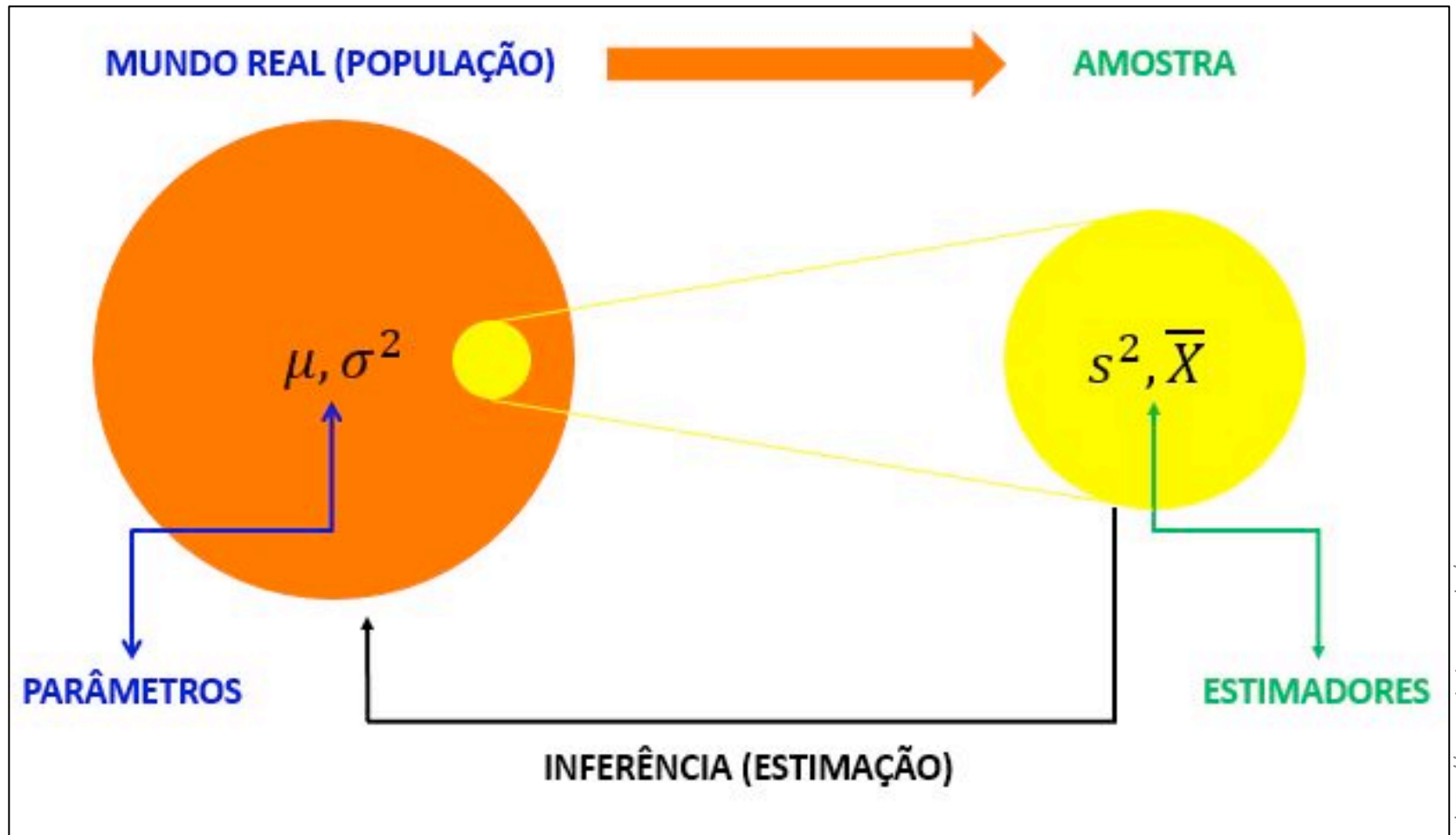
INCERTEZAS SOBRE A QUALIDADE  
DA ESTIMATIVA DOS  
PARÂMETROS POPULACIONAIS



**AMOSTRAL**



# DIFERENÇA ENTRE PARÂMETROS E ESTIMADORES



<http://oestatistico.com.br/>

Desconhecidos,  
Mas **FIXOS**

**Importância de definir  
claramente sua população**

Utilizam os valores da  
amostra tomada

Estimativa pontual ou  
intervalos

## Quais as características de um BOM ESTIMADOR?

**SEM VIÉS** - Média de várias medidas deve coincidir com o parâmetro original

**CONSISTENTE** - Maior tamanho da amostra, o valor converge para o parâmetro original e a variância diminui

**EFICIENTE** - Dado um mesmo tamanho de amostra, o estimador mais eficiente será o que apresentar menor variância

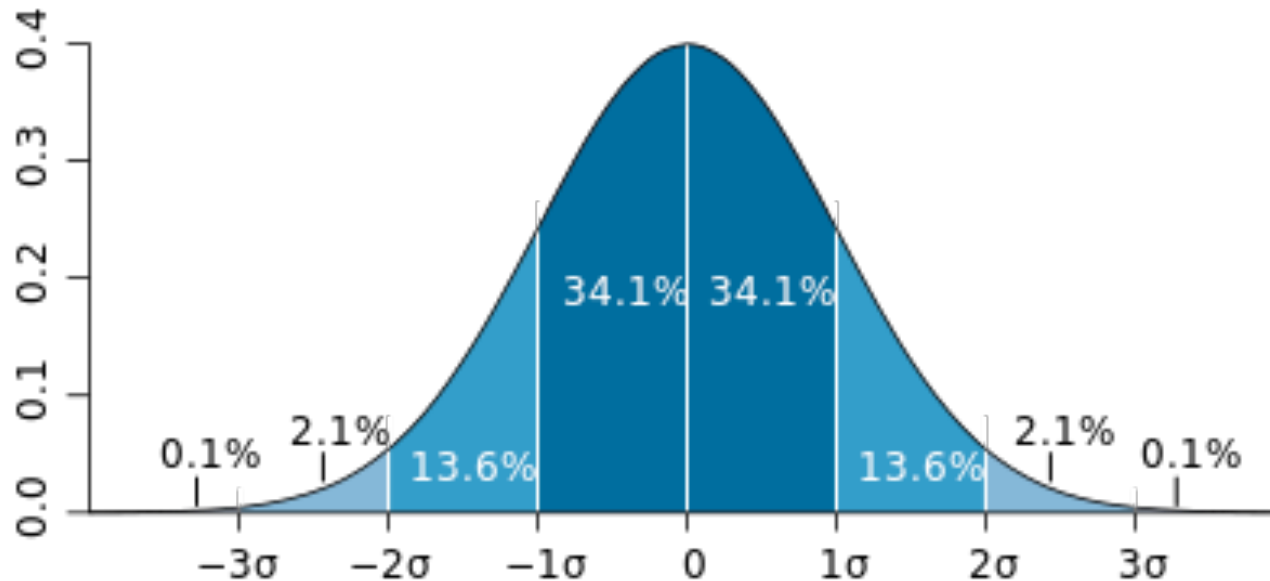
DEFINIÇÕES BASEADAS EM PROCEDIMENTOS MATEMÁTICOS E/OU COMPUTACIONAIS

NÃO SE PREOCUPEM COM ISSO!!

EXISTE TODA UMA ÁREA DA ESTATÍSTICA VOLTADA PARA ENCONTRAR BONS ESTIMADORES

# DIFERENTES DISTRIBUIÇÕES DE PROBABILIDADES TÊM DIFERENTES ESTIMADORES

Variável descrita por uma Distribuição Normal



Média e Variância da AMOSTRA  
são **bons estimadores** da  
Média e Variância da POPULAÇÃO

Mediana não é tão eficiente quanto a Média para a distribuição normal

# ALGUNS PARÂMETROS E SEUS ESTIMADORES PONTUAIS

Parameter	Statistic	Formula
Mean ( $\mu$ )	$\bar{y}$	$\frac{\sum_{i=1}^n y_i}{n}$
Median	Sample median	$y_{(n+1)/2}$ if $n$ odd $(y_{n/2} + y_{(n/2)+1})/2$ if $n$ even
Variance ( $\sigma^2$ )	$s^2$	$\sum_{i=1}^n \frac{(y_i - \bar{y})^2}{n-1}$
Standard deviation ( $\sigma$ )	$s$	$\sqrt{\sum_{i=1}^n \frac{(y_i - \bar{y})^2}{n-1}}$
Median absolute deviation (MAD)	Sample MAD	$\text{median}[ y_i - \text{median} ]$
Coefficient of variation (CV)	Sample CV	$\frac{s}{\bar{y}} \times 100$
Standard error of $\bar{y}$ ( $\sigma_{\bar{y}}$ )	$s_{\bar{y}}$	$\frac{s}{\sqrt{n}}$
95% confidence interval for $\mu$		$\bar{y} - t_{0.05(n-1)} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{y} + t_{0.05(n-1)} \frac{s}{\sqrt{n}}$

Diferenças entre médias ( $\mu_1 - \mu_2$ ) pode ser um **parâmetro** (exemplo do Manguezal)

Amplitude ("*range*") dos valores da amostra não é um bom parâmetro

PARA UMA DADA AMOSTRA, UM ESTIMADOR VAI PRODUZIR UM **VALOR** (OU INTERVALO)



**ESTIMATIVA**



P.ex.: pontuação média

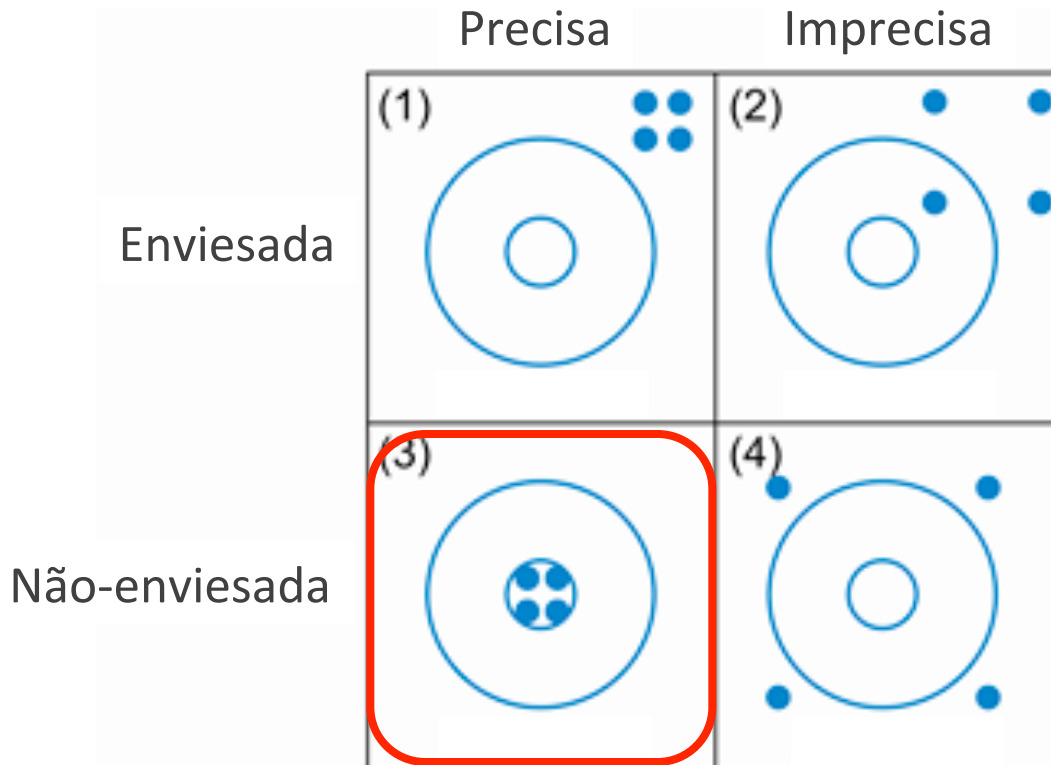


**NOSSO OBJETIVO:** Obter uma **estimativa** confiável com o menor esforço amostral

# O QUE É UMA ESTIMATIVA CONFIÁVEL?

**SEM VIÉS** - O valor obtido com a amostra estatística deve ser igual ao parâmetro. Não deve subestimar ou superestimar o parâmetro populacional

**PRECISA** - A maior parte dos valores obtidos na amostra, deve estar nas proximidades do parâmetro populacional (ERROS PEQUENOS)



**Trabalhando um exemplo hipotético...**

**Em uma primeira fase estamos interessados em modelar as respostas reprodutivas de uma espécie rara de planta, visando assegurar sua manutenção a longo prazo e explorar um recurso**

### **Algumas informações prévias:**

- A espécie ocorre nas partes **mais altas** de uma cadeia de montanhas
- A espécie é geograficamente rara, mas localmente abundante
- É uma espécie típica de sub-bosque
- Possui um fármaco de importância comercial extraído das sementes
- Toda a população produz sementes em um mesmo período
- Ainda não sabemos quase nada dessa espécie

### **Objetivo inicial:**

**Estimar o número médio de sementes produzidas por essa espécie**



Vocês têm recursos para fazer 10 unidades amostrais (parcelas de 50 x 50 m).

Considerando a paisagem abaixo e as características da planta indicadas anteriormente, indiquem na figura 1, usando o símbolo "x", como alocariam as UAs.

Oeste (W)



Oceano Atlântico - Leste (E)

### **Atividade A)**

Expliquem, no verso da folha, que motivos levaram vocês a distribuir as 10 unidades "x" dessa forma

### **Atividade B)**

Anotem, no verso da folha, quais aspectos da paisagem representada nesta figura vocês acreditam que poderiam interferir na produção de sementes da espécie

## Parabéns!!! Você acaba de receber um financiamento!!

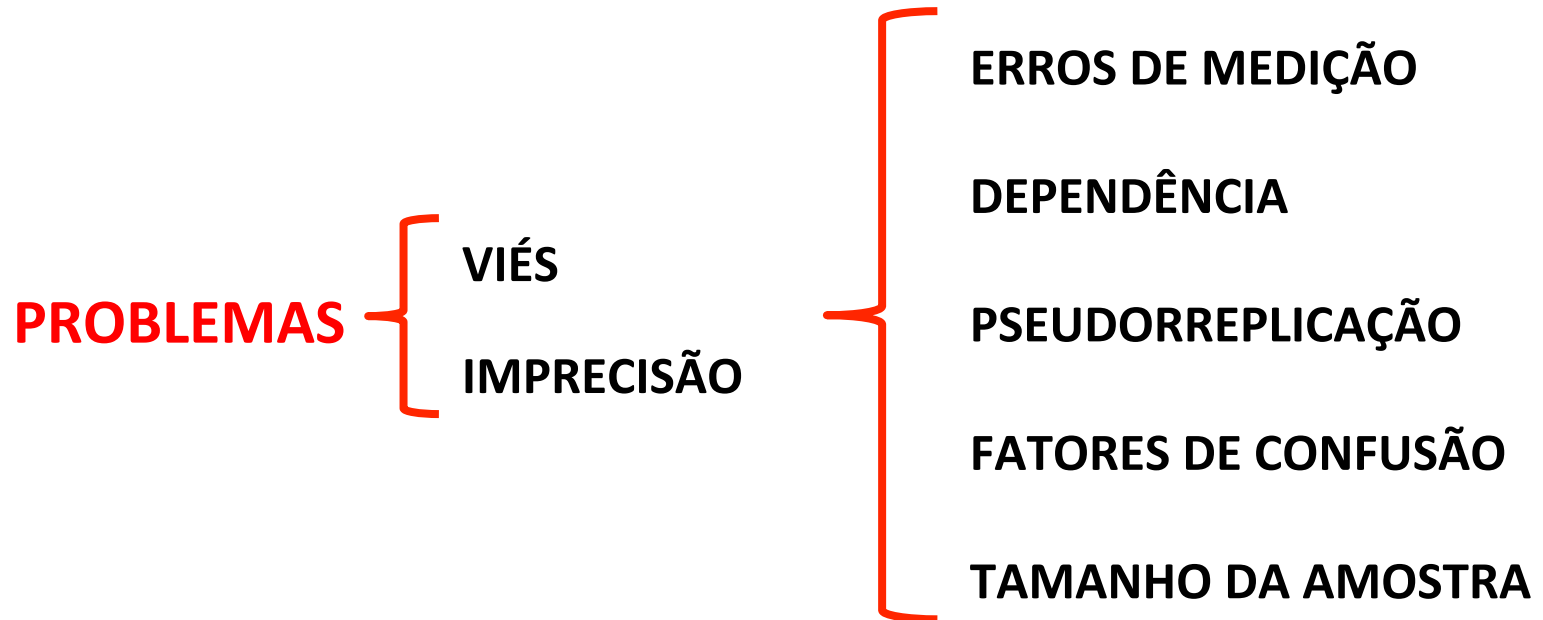
Vocês têm recursos para fazer 60 unidades amostrais (parcelas de 20 x 20m) .  
Indiquem na figura 1, com o símbolo "o", como alocariam essas UAs.



## **Atividade C)**

Expliquem, no verso da folha, que motivos levaram vocês a distribuir as 60 unidades "o" dessa forma

# DIFICULDADES PARA SE OBTER UMA BOA ESTIMATIVA



# A IMPORTÂNCIA DESSES PROBLEMAS ESTÁ RELACIONADA À **FORÇA DE INFERÊNCIA** DESEJADA

## NOSSO OBJETIVO

PERGUNTA CIENTÍFICA - HIPÓTESE - DELINEAMENTO - COLETA - ANÁLISE - CONCLUSÃO



## **FORÇA DE INFERÊNCIA**

**Relação entre validade interna/externa**

# TIPOS DE ESTUDOS ECOLÓGICOS

**OBSERVACIONAIS** - Variáveis **preditoras** não controladas

DESCRITIVOS

ANALÍTICOS

CONTROLE - IMPACTO

**EXPERIMENTAIS** - Variáveis **preditoras** controladas



LABORATÓRIO  
CAMPO

Adaptado de Manly (1992), Schwarz (1998) e Eberhardt & Thomas (1991)

**A nomenclatura e a classificação dos diferentes estudos variam muito entre autores!**

## O CONCEITO DE VALIDADE

### **Validade interna (causalidade)**

Varição observada na variável resposta é claramente causada pela variável preditora que está sendo testada



### **Validade externa (generalidade)**

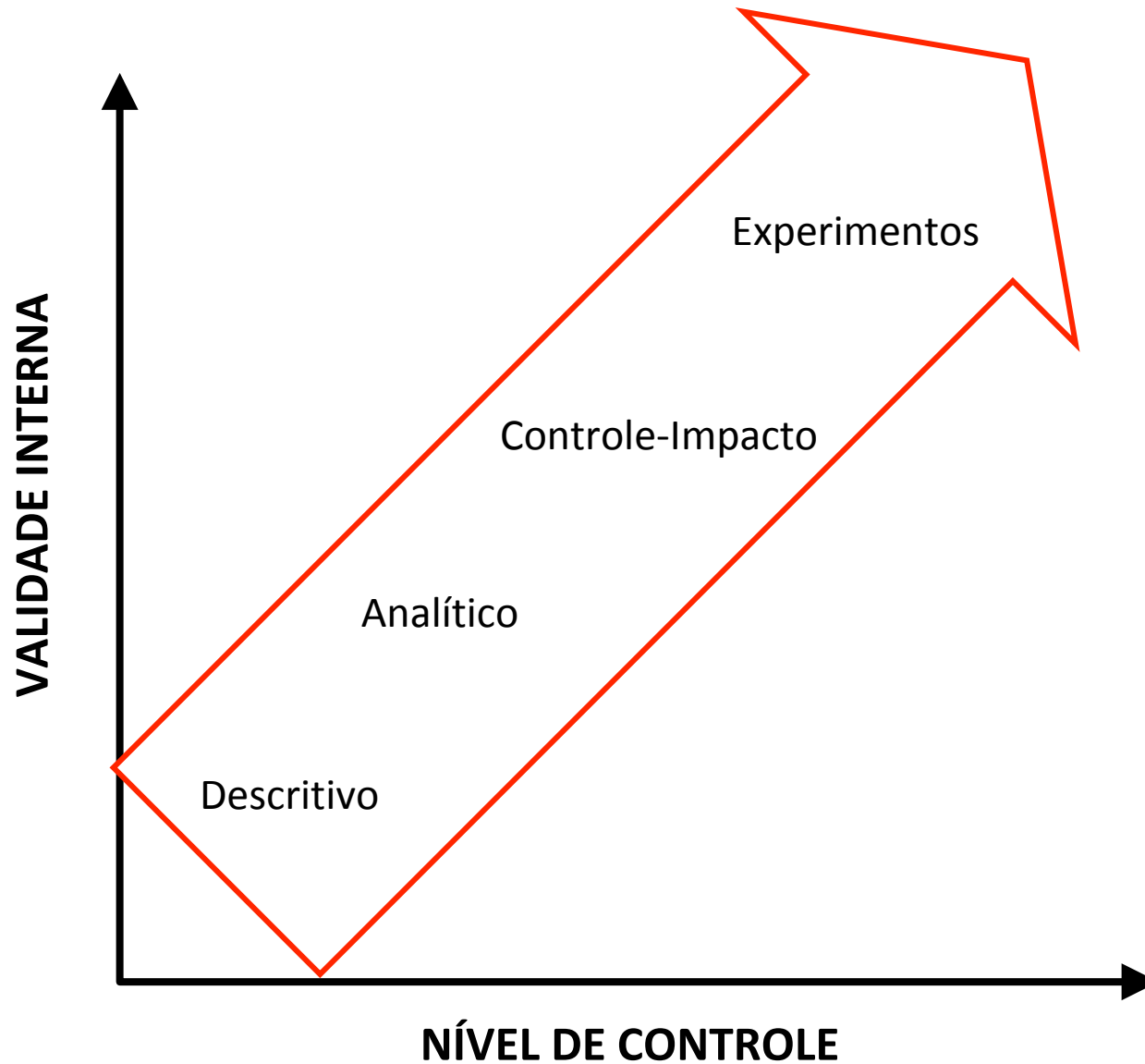
A associação (ou estimativa) observada pode ser generalizada para a população toda, ou para outros grupos, outros contextos



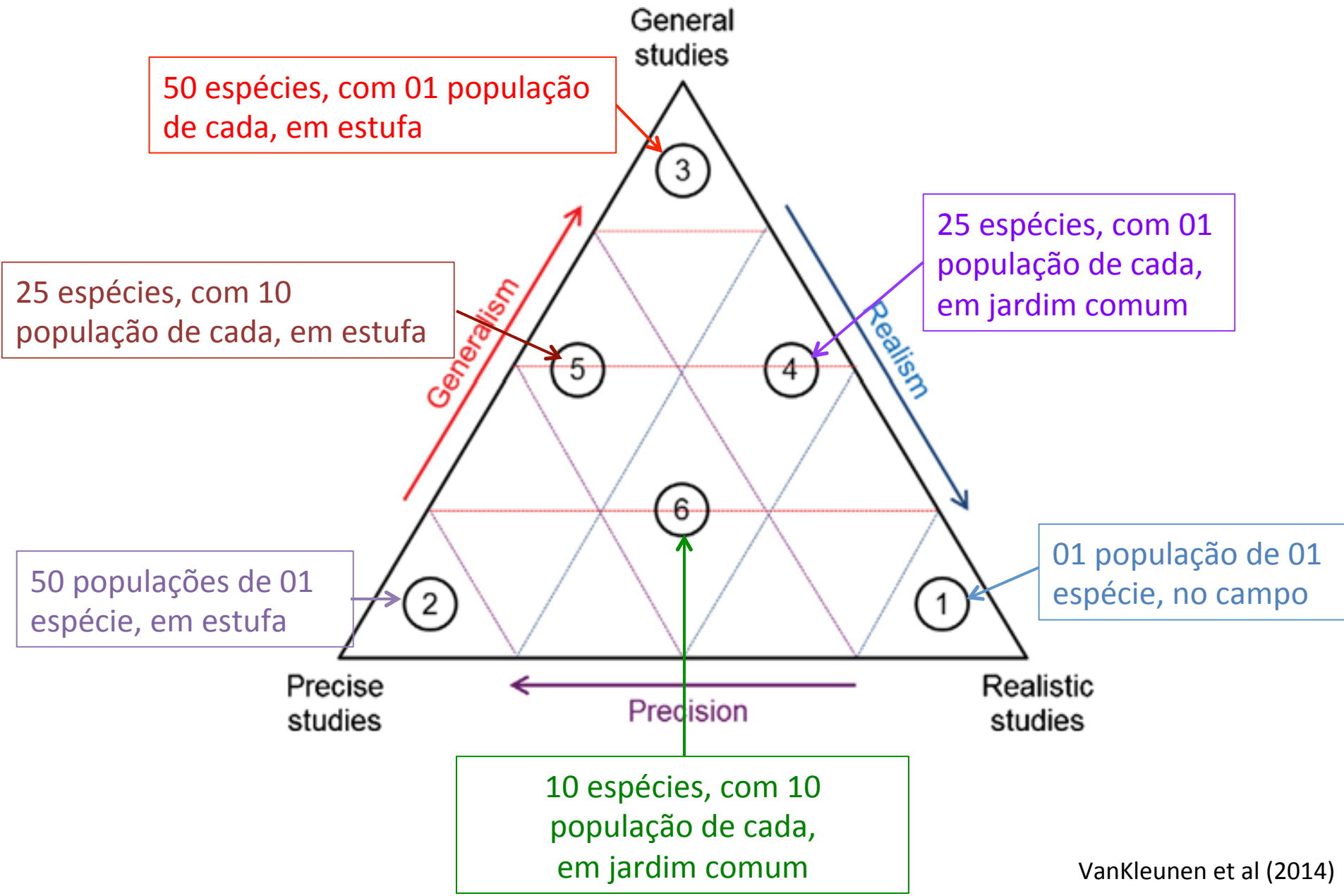
**Alguns autores sugerem que esses atributos seriam excludentes**



Maior nível de **controle das preditoras** -> maior **VALIDADE INTERNA**



# Triângulo de Levins



**LEMBRAR SEMPRE:**

**QUAL É A SUA PERGUNTA CIENTÍFICA?**

# **IDENTIFICANDO OS PRINCIPAIS PROBLEMAS**

# DEPENDÊNCIA

O valor de uma dada unidade amostral é influenciado por outra unidade amostral

- **DEPENDÊNCIA ESPACIAL**

- **DEPENDÊNCIA TEMPORAL**

- **DEPENDÊNCIA DE ORIGEM** (p. ex.: sub-amostras de um mesmo organismo)



**DEPENDÊNCIA POSITIVA:**  
Unidades amostrais dependentes  
são mais similares entre si  
**Padrão agregado**  
**Muito comum em Ecologia**



**DEPENDÊNCIA NEGATIVA:**  
Unidades amostrais dependentes  
são mais diferentes entre si  
**Ex. Alelopatia**

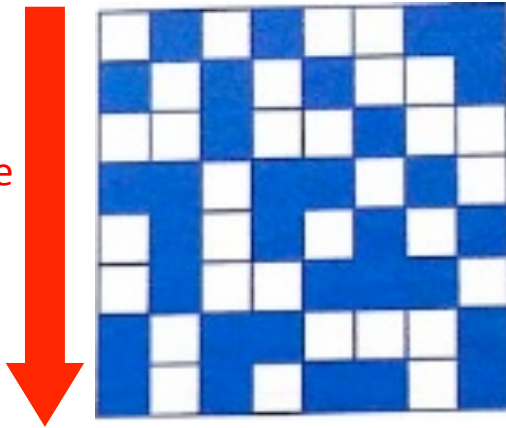
# Exemplo com **DEPENDÊNCIA ESPACIAL POSITIVA**

## COMPARANDO MÉDIAS DE DUAS AMOSTRAS

AMOSTRA 1

(unidades amostrais independentes)

umidade  
do solo

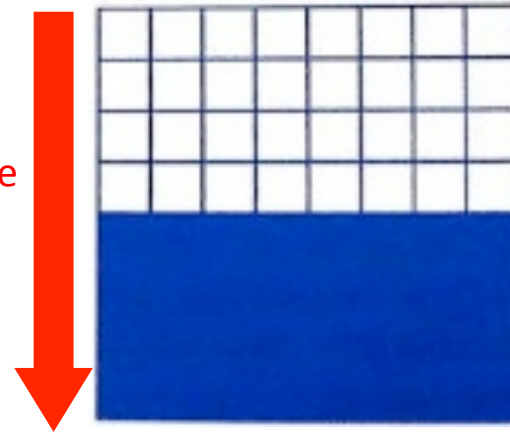


ESTIMATIVA não  
enviesada e com  
maior variação

AMOSTRA 2

(unidades amostrais dependentes)

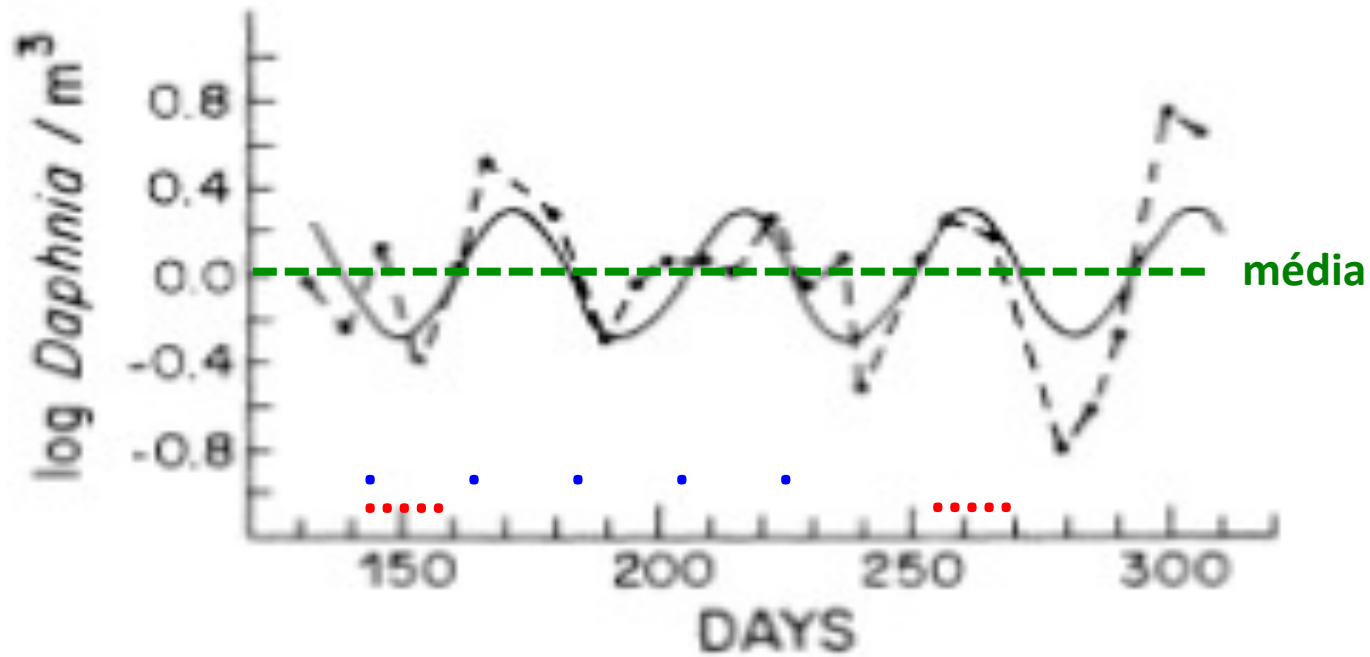
umidade  
do solo



ESTIMATIVA  
enviesada e com  
menor variação

**Maior chance de ERRO TIPO I -> HIPÓTESE NULA REJEITADA ERRONEAMENTE**

As mesmas ideias também se aplicam à **DEPENDÊNCIA TEMPORAL**



Quanto mais próximas as datas de coleta, mais similares serão os valores -> menor variação

Maior espalhamento -> maior variação

Qual estima melhor a média?

# DEPENDÊNCIA

## COMO EVITAR:

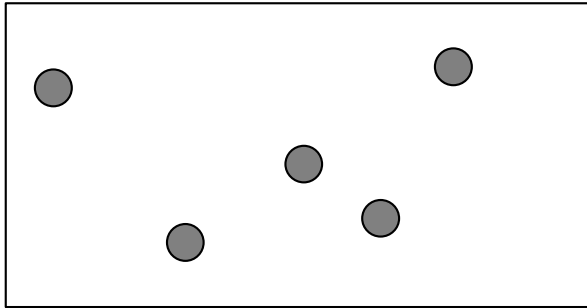
- Conhecimento prévio do sistema
- Aumentando a distância/tempo entre as unidades amostrais (UA)
- Aumentando a heterogeneidade de distâncias/tempos entre as UAs
- Aleatorização (+ tamanho amostral adequado)
- Tomando amostras diferentes a cada tempo
- Ver adiante exemplos de delineamentos



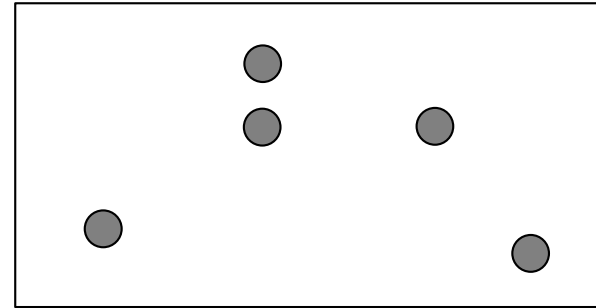
# PSEUDORREPLICAÇÃO

Unidades amostrais tomadas em uma escala e **inferência** feita em outra escala  
Não são réplicas verdadeiras (são sub-amostras) e **não são independentes**

Área queimada



Área não queimada



Círculo cinza = ponto de amostragem de biomassa de micro-organismos de solo

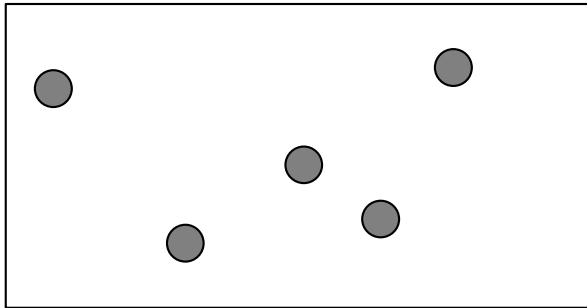
Qual(is) pergunta(s) poderia(m) ser respondida(s) com os dados obtidos nesse estudo?

Quais são as unidades amostrais no exemplo acima?

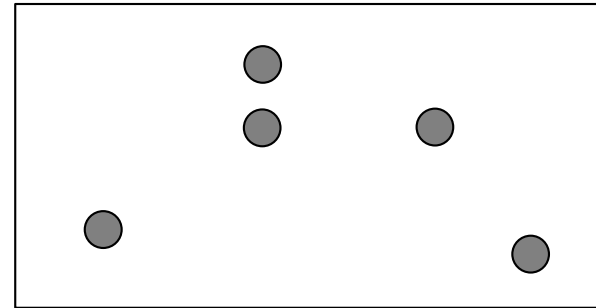
## SERIA UMA PSEUDORREPLICAÇÃO SE A PERGUNTA FOSSE:

*"Áreas queimadas têm menos biomassa de micro-organismos no solo?"*

Área queimada



Área não queimada



Círculo cinza = ponto de amostragem de biomassa de micro-organismos de solo

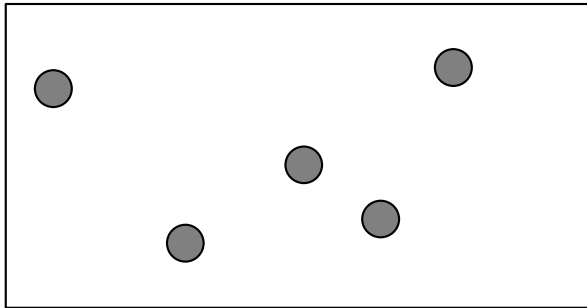
As **unidades amostrais** seriam os **quadrados** (**n=1** em cada condição) e os **círculos** são sub-amostras

**BAIXA PRECISÃO E BAIXA VALIDADE EXTERNA**

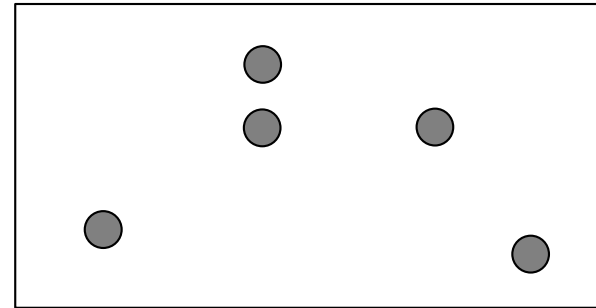
## NÃO SERIA UMA PSEUDORREPLICAÇÃO SE A PERGUNTA FOSSE:

*"Essa área queimada tem menos biomassa de micro-organismos no solo do que essa área não queimada?"*

Área queimada



Área não queimada



Círculo cinza = ponto de amostragem de biomassa de micro-organismos de solo

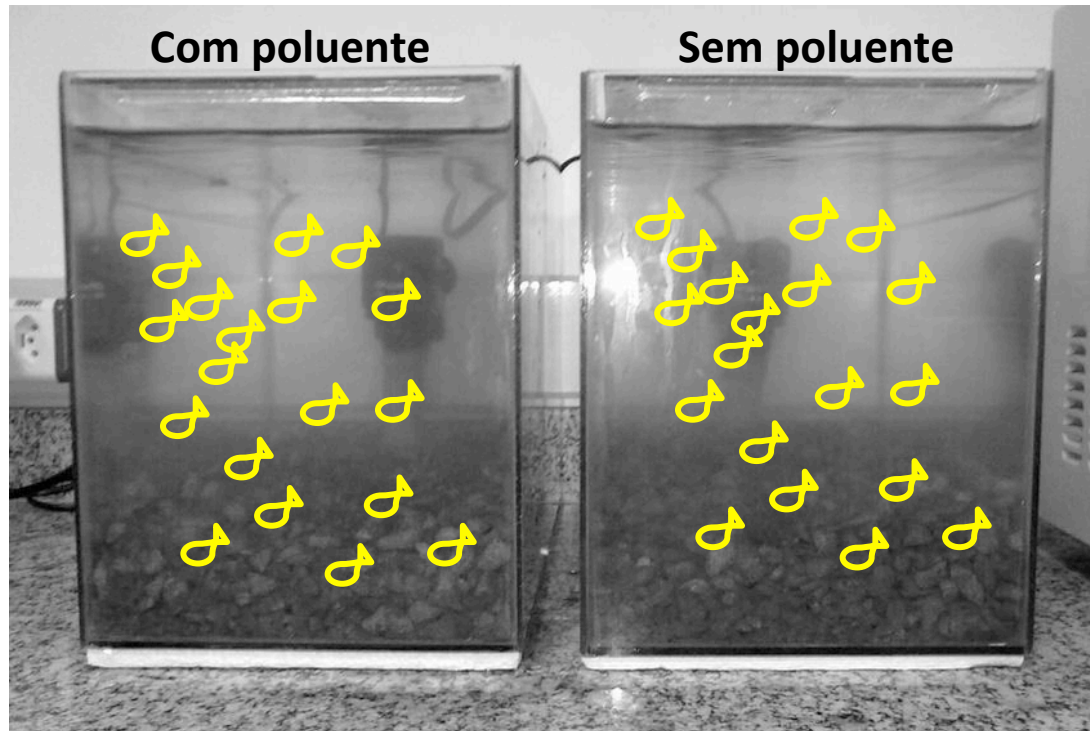
Nesse caso: as **unidades amostrais** seriam os **círculos** (n=5 em cada área)

**"BOA" PRECISÃO, MAS BAIXA VALIDADE EXTERNA**

# PSEUDORREPLICAÇÃO

## Outro exemplo muito comum

20 Peixes em cada aquário



Quais são as unidades amostrais no exemplo acima?

### Possível solução:

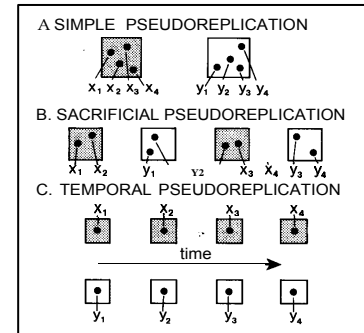
Réplicas temporais dos mesmos aquários (substituindo os peixes e aleatorizando sempre qual aquário recebe ou não o poluente)

# PSEUDORREPLICAÇÃO (tópico polêmico)

*Ecological Monographs* 54(2), 1984, pp. 187-211  
© 1984 by the Ecological Society of America

## PSEUDOREPLICATION AND THE DESIGN OF ECOLOGICAL FIELD EXPERIMENTS

STUART H. HURLBERT  
Department of Biology, San Diego State University,  
San Diego, California 92182 USA



### *On misinterpretations of pseudoreplication and related matters: a reply to Oksanen*

Stuart H. Hurlbert, Stuart H. Hurlbert, Dept of Biology and Center for Inland Waters, San Diego State Univ., San Diego, California 92182, USA. (shurlbert@sunstroke.sdsu.edu)

2004, a luta continua...

Sub-amostras aumentam a **precisão** de uma estimativa

MAS não podem ser tratadas como réplicas independentes! -> **Correção analítica**

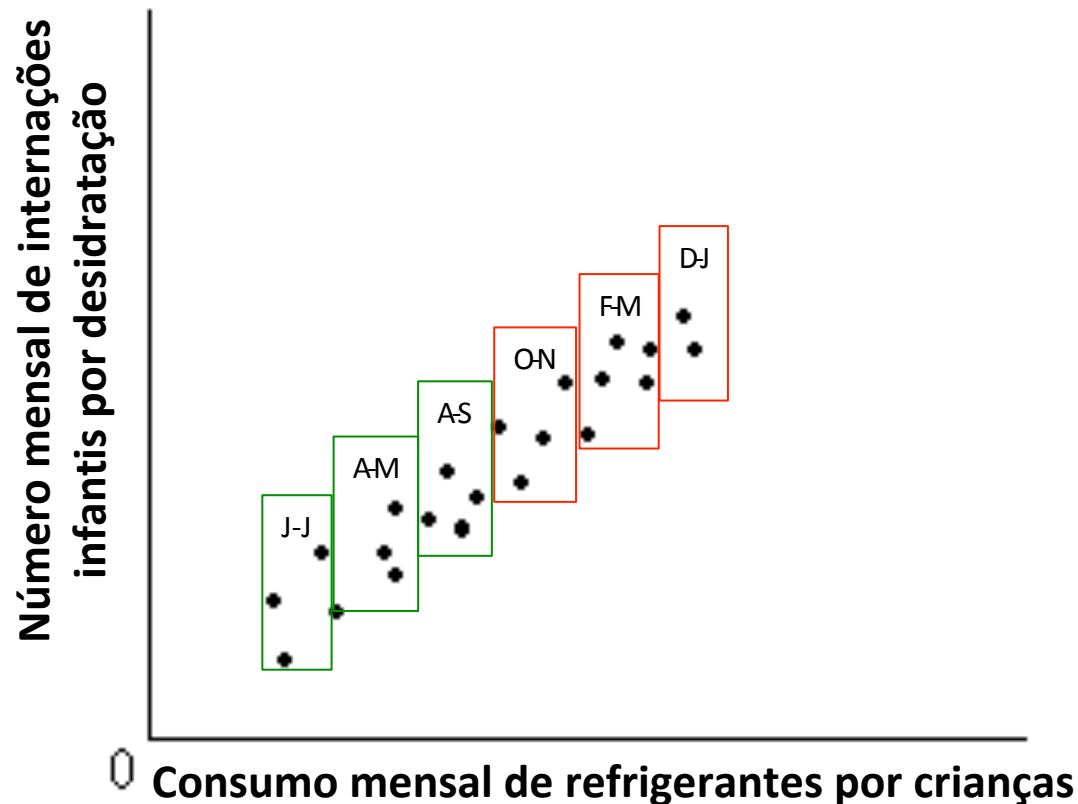
Em geral, é melhor investir em  
mais unidades amostrais (réplicas)  
do que em sub-amostras

X

Mas, existem situações que é  
importante investir em sub-  
amostras para não perder réplicas

# FATORES DE CONFUSÃO

A relação entre duas variáveis pode ser explicada por outro fator

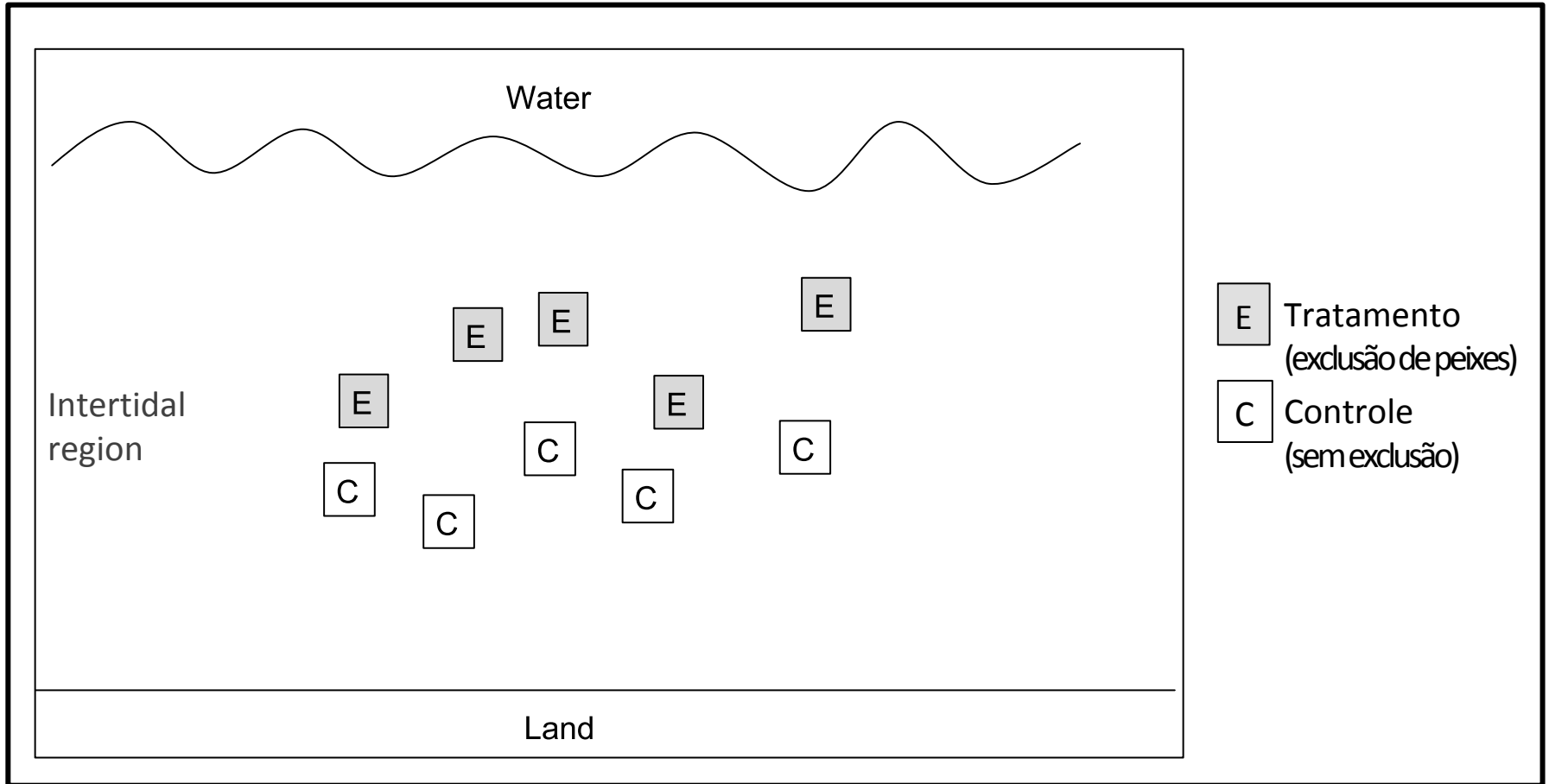


A variável "mês de amostragem" (ou "estação do ano") não foi incluída na análise

# FATORES DE CONFUSÃO

O efeito de um tratamento pode ser explicado por outro fator

Estudo sobre predação da comunidade bentônica por peixes na zona entre-marés



Nesse exemplo, a maior proximidade com o mar poderia explicar as diferenças entre E e C

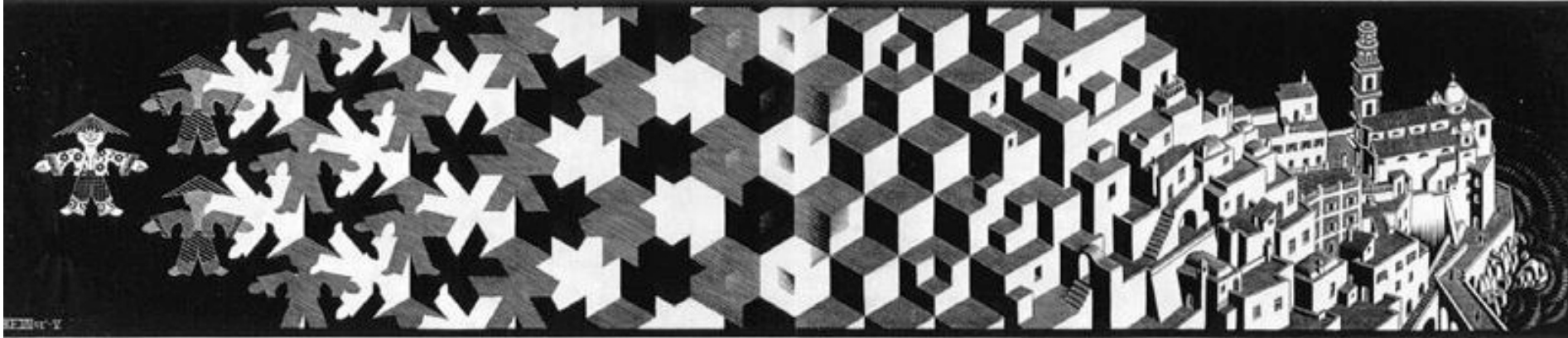
# FATORES DE CONFUSÃO

## COMO EVITAR:

- Conhecimento prévio e/ou estudo-piloto
- Manipulação ou controle de condições
- Medição de variáveis adicionais (Covariáveis) - Antes/Durante/Depois



# FATORES DE CONFUSÃO



Descrição -> Previsões -> Processos -> Mecanismos

Descritivo

Controle-Impacto

Controle-Impacto

**Experimentos**

Analítico

Analítico

**Experimentos**

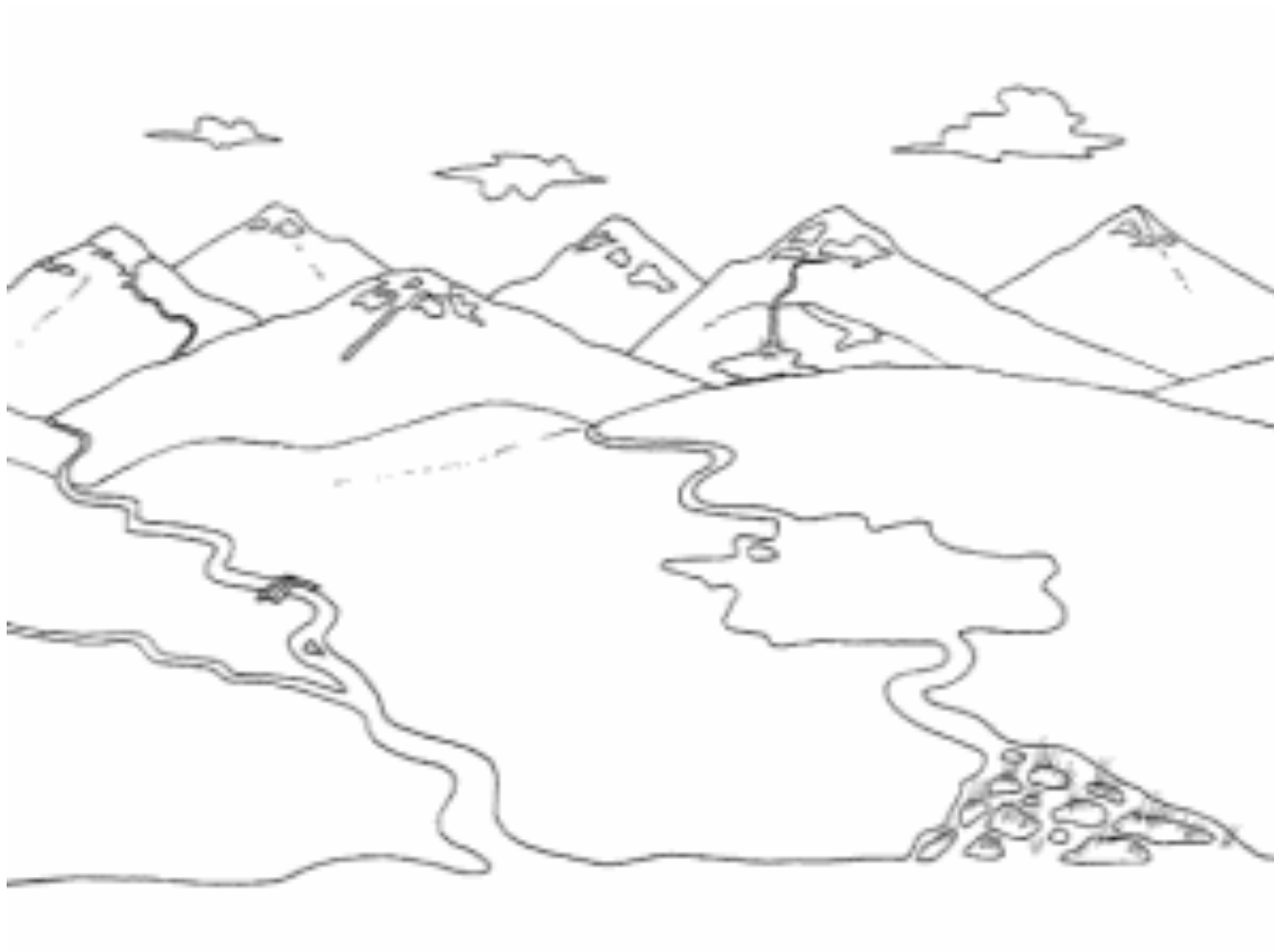
**Experimentos**



Vocês identificam potenciais **problemas** na amostragem anterior que realizaram?

Oeste (W)

Sul (S)



Norte (N)

Oceano Atlântico - Leste (E)

**Gostariam de realocar suas unidades amostrais? Pensem sobre isso...**

# **DELINEAMENTOS PARA ESTUDOS OBSERVACIONAIS** **(não exclusivamente)**

# AMOSTRAGEM ALEATÓRIA SIMPLES

## PRINCÍPIO BÁSICO:

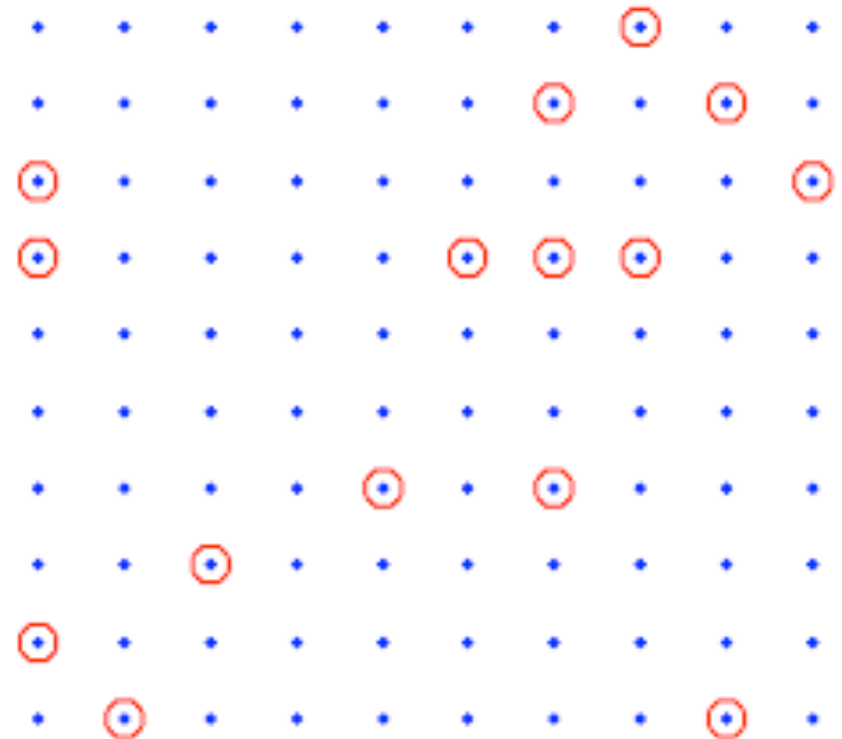
Qualquer unidade amostral deve ter a mesma probabilidade de ser amostrada

### PROCEDIMENTOS:

DEFINIR CLARAMENTE A **POPULAÇÃO (\*)**  
(limites espaciais e temporais )

DEFINIR **UNIDADES AMOSTRAIS(\*)**

DEFINIR A **FORMA DE ALEATORIZAÇÃO (\*)**



(\*) importante para todos os próximos delineamentos

# ALEATORIZAÇÃO

Quais métodos conhecem para definir onde estabelecer uma unidade amostral?

Muitos métodos inadequados (não atendem as premissas da aleatoriedade):

- Ir contando e pedir para alguém dizer para parar
- Atirar pedras ou outros objetos para trás

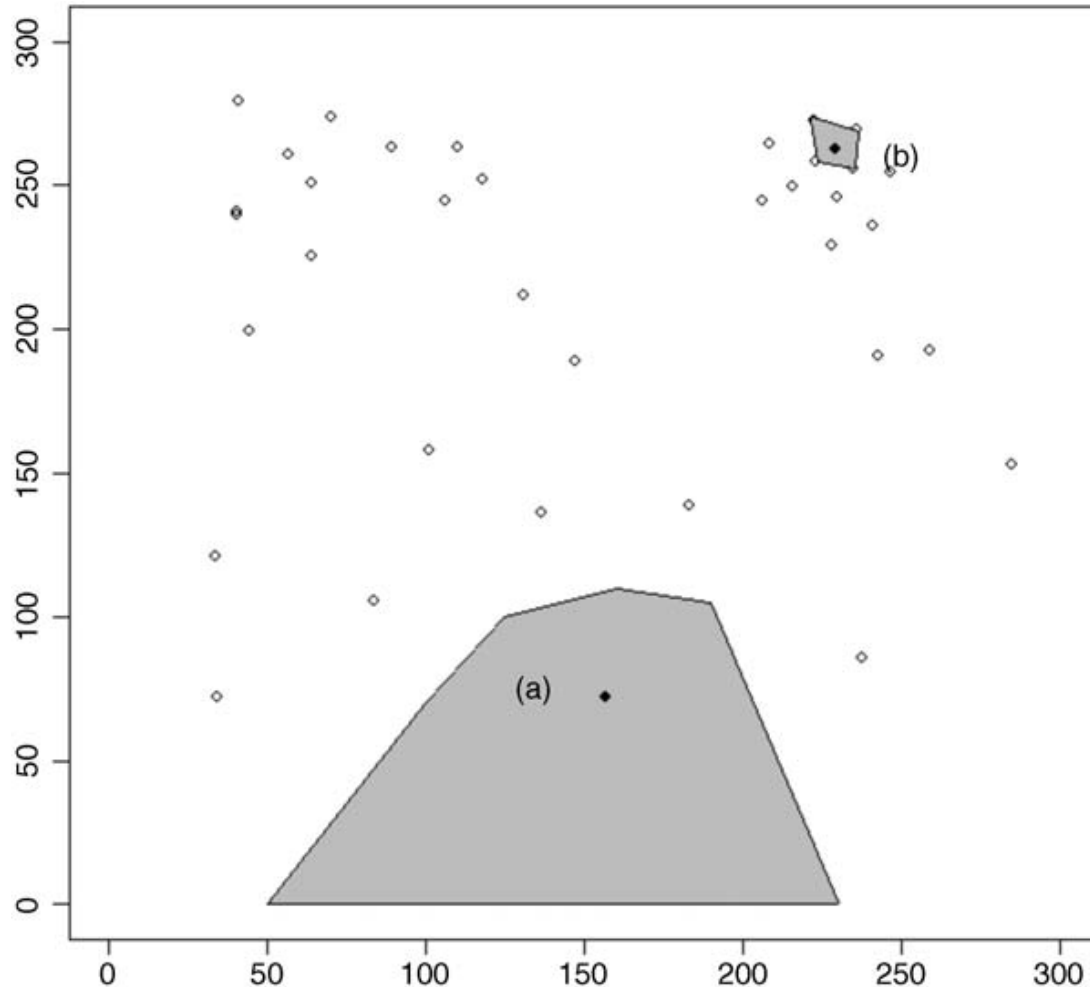
*"Some investigators locate plots by **throwing a rock over your shoulder**, or **walking a certain number of steps with their eyes closed**. Not only are such techniques **dangerous**, they also do not represent random sampling. We call such techniques 'haphazard'." Michael Palmer*



## Outro método inadequado (apesar de muito usado):

- Sortear um ponto xy e amostrar indivíduo mais próximo

**As árvores (a) e (b) dessa parcela têm a mesma chance de serem amostradas por esse método?**



As áreas em cinza definem todos os potenciais pontos xy que indicariam aquela árvore para ser amostrada

# ALEATORIZAÇÃO

## Alguns métodos adequados:

- Tabelas de números aleatórios
- Números aleatórios gerados por programas de computador (ressalvas)
- Últimos dígitos de cronômetro (milissegundos)

20	17	42	01	72	33	94	55	89	65	58	60
74	49	04	27	56	49	11	63	77	79	90	31
94	70	49	49	05	74	64	00	26	07	23	00
22	15	78	49	74	37	50	94	13	90	08	14
93	29	12	20	26	22	66	98	37	53	82	62
45	04	77	48	87	77	66	91	42	98	17	26
44	91	99	08	72	87	33	58	12	08	91	12
16	23	91	95	97	98	52	49	40	37	21	46
04	50	65	37	99	57	74	98	93	99	78	30
32	70	17	05	79	58	50	26	54	30	01	88
03	64	59	55	85	63	49	46	61	89	33	79
62	49	00	67	28	96	19	65	13	44	78	39
61	00	95	85	86	94	64	17	47	67	87	59
89	03	90	40	10	60	18	43	97	37	68	97



# AMOSTRAGEM ALEATÓRIA SIMPLES

## VANTAGENS:

SE FOI POSSÍVEL SEGUIR TODOS OS PROCEDIMENTOS INDICADOS :

$$\begin{array}{l} \text{MÉDIA AMOSTRAL} \\ \text{(Estimador)} \end{array} \quad \frac{\sum_{i=1}^n y_i}{n} = \begin{array}{l} \text{MÉDIA POPULACIONAL} \\ \text{(Parâmetro)} \end{array}$$

$$\begin{array}{l} \text{VARIÂNCIA AMOSTRAL} \\ \text{(Estimador)} \end{array} \quad \sum_{i=1}^n \frac{(y_i - \bar{y})^2}{n-1} = \begin{array}{l} \text{VARIÂNCIA POPULACIONAL} \\ \text{(Parâmetro)} \end{array}$$

(n-1) é usado ao invés de n para corrigir o estimador

**TESTES DE HIPÓTESES PODEM SER REALIZADOS COM SEGURANÇA**



# AMOSTRAGEM ALEATÓRIA SIMPLES

## DESVANTAGENS E PROBLEMAS:

- REQUER UM TAMANHO GRANDE DE AMOSTRA
- POSSÍVEL DIFICULDADE DE ACESSO AOS PONTOS DEFINIDOS
- MUITO ESFORÇO QUANDO OBJETOS ESTÃO AGREGADOS OU SÃO RAROS (localmente)
- SE HOUVER HETEROGENEIDADE, AMBIENTES RAROS PODEM NÃO SER AMOSTRADOS

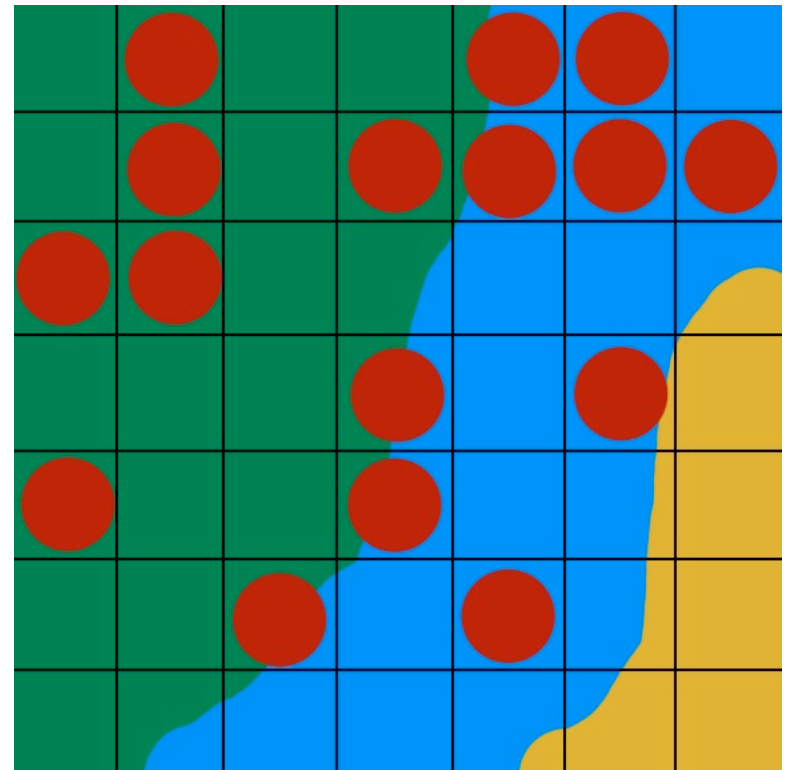
### Amostragem aleatória

• Procedimento estatístico excelente

• Pode exigir muito tempo para encontrar locais

Pode pular alguns habitats

Muitos locais podem estar vazios



# AMOSTRAGEM ALEATÓRIA SIMPLES

## SOLUÇÕES PARA OS PROBLEMAS:

- ESTUDO PILOTO E BUSCA POR CONHECIMENTOS PRÉVIOS
- TÉCNICAS ANALÍTICAS POSTERIORES (Procedimentos de Monte Carlo)
- MEDIÇÕES ADICIONAIS (PARA EVITAR FATORES DE CONFUSÃO; PÓS-ESTRATIFICAÇÃO)
- OPÇÃO POR OUTROS ESQUEMAS DE AMOSTRAGEM (ver a seguir)

# AMOSTRAGEM ALEATÓRIA ESTRATIFICADA

## PRINCÍPIO BÁSICO:

Divide a população em "estratos" e aloca unidades amostrais dentro de cada estrato

## PROCEDIMENTOS:

### DEFINIR ESTRATOS

Princípio: internamente mais homogêneos que entre

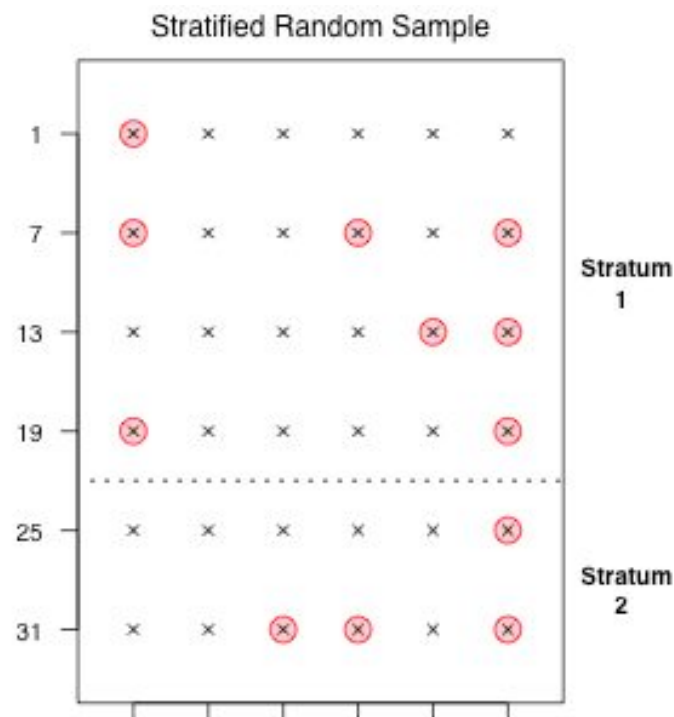
### QUANTIFICAR OS ESTRATOS

PRÉ: Se os fatores da estratificação já forem conhecidos

PÓS: Se os fatores não forem conhecidos *a priori*

### AMOSTRAR OS ESTRATOS

Dentro de cada estrato, a amostragem deve ser **aleatória**



Mas, precisa manter um número mínimo de réplicas em cada estrato (Gotelli & Ellison  $\approx 10$ )



# AMOSTRAGEM ALEATÓRIA ESTRATIFICADA

## DESVANTAGENS E PROBLEMAS:

- NO CASO DE HETEROGENEIDADE PERCEPTÍVEL

Necessidade de conhecimento prévio e/ou de tomada de dados adicionais

- NO CASO DE HETEROGENEIDADE IMPERCEPTÍVEL

Erro no "grão" para definir estratos

## SOLUÇÕES:

- ANALÍTICAS (**PÓS-ESTRATIFICAÇÃO** com dados adicionais)

- MEDIÇÕES ADICIONAIS

- OPÇÃO POR OUTROS ESQUEMAS DE AMOSTRAGEM

# AMOSTRAGEM SISTEMÁTICA

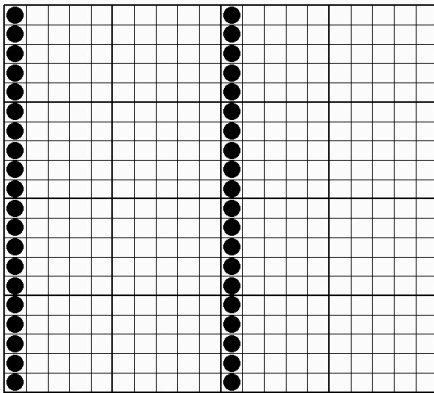
## PRINCÍPIO BÁSICO:

A partir de um ponto inicial, as unidades amostrais estão espalhadas a uma distância fixa

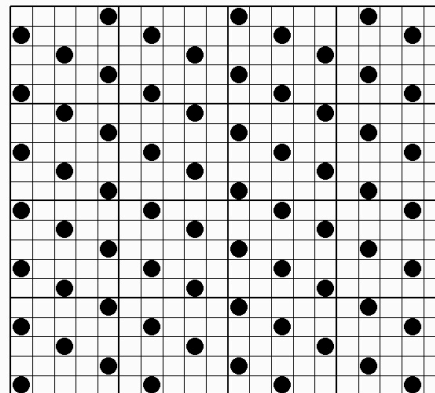
## PROCEDIMENTO FUNDAMENTAL:

- DEFINIR O ESPAÇAMENTO

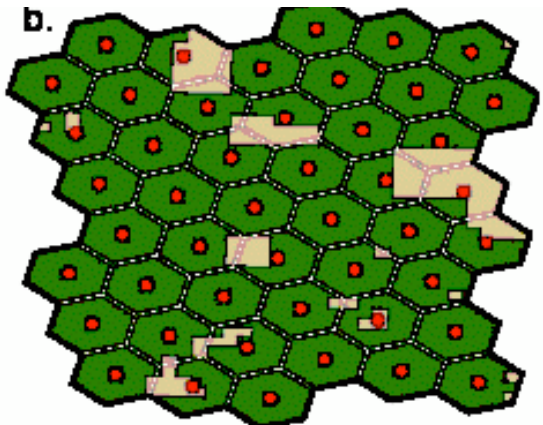
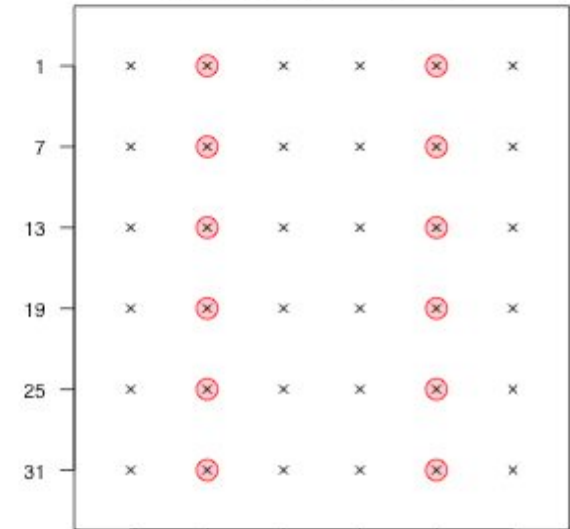
SYSTEMATIC SAMPLE



SYSTEMATIC SAMPLE



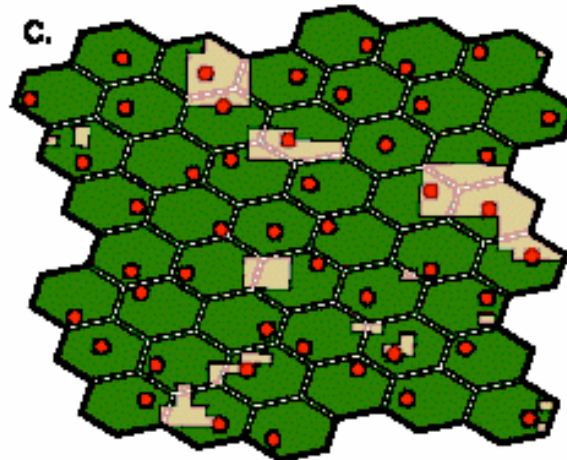
Systematic Random Sample



# AMOSTRAGEM SISTEMÁTICA

## VANTAGENS:

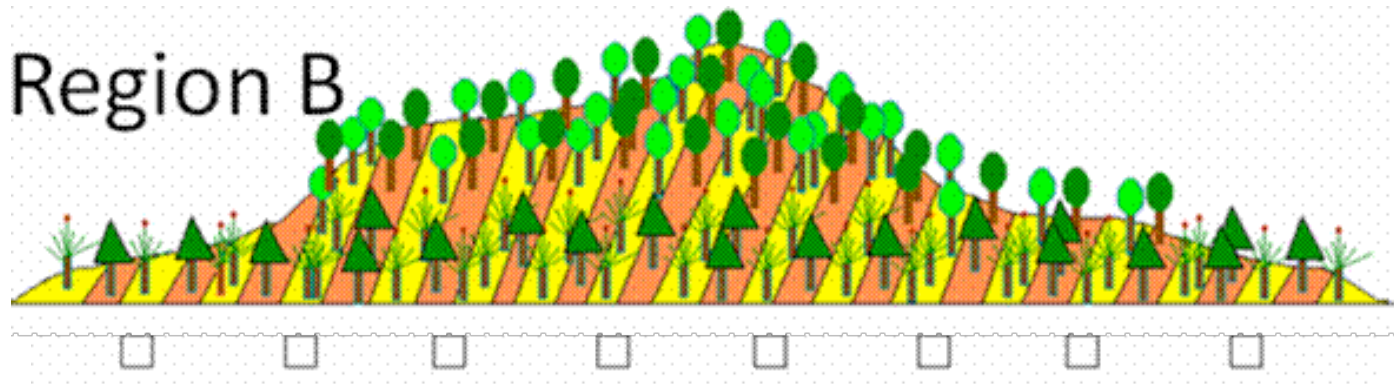
- FÁCIL EXECUÇÃO
- ACESSO FACILITADO AOS PONTOS DE MEDIÇÃO/ FÁCIL RECUPERAÇÃO PARA RECENSOS
- ADEQUADA PARA REVELAR MUDANÇAS AO LONGO DE GRADIENTES
- ESPALHA MELHOR AS UNIDADES AMOSTRAIS (**INTERSPERSÃO**) - Maior representatividade
- PODE SER **ASSOCIADA** COM OUTROS ESQUEMAS DE AMOSTRAGEM



# AMOSTRAGEM SISTEMÁTICA

## DESVANTAGENS E PROBLEMAS:

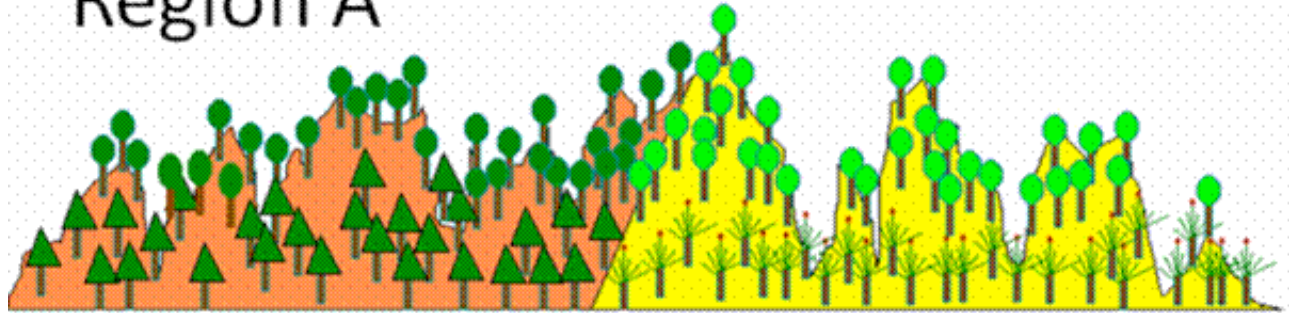
- DIFICULDADE DE DEFINIR O GRAU DE ESPAÇAMENTO
- AUMENTA MUITO A CHANCE DE HAVER **DEPENDÊNCIA** ENTRE UNIDADES AMOSTRAIS
- SINCRONIA INDESEJADA COM ALGUM FATOR AMBIENTAL (obscurece a variância real)





# ESCALA (GRÃO E EXTENSÃO)

Region A



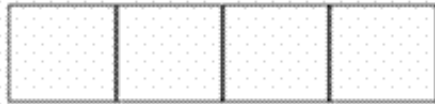
Grão Pequeno X Extensão Pequena



Grão Pequeno X Extensão Grande



Grão Grande X Extensão Pequena



Grão Grande X Extensão Grande



Region B



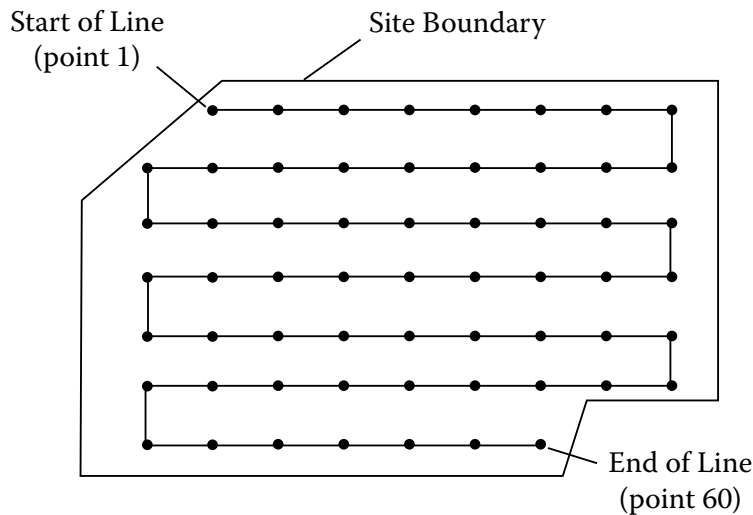
Qual seria a melhor e qual seria a pior amostragem para cada uma das regiões?

# AMOSTRAGEM SISTEMÁTICA

## SOLUÇÕES:

- MEDIÇÕES ADICIONAIS ANTERIORES (Evitar sincronia; definir espaçamento)
- ASSOCIAR COM OUTROS ESQUEMAS DE AMOSTRAGENS
- ANALÍTICAS (*LAG ANALYSIS*, PÓS-ESTRATIFICAÇÃO; SERPENTINA)

Existem formas analíticas de incorporar o efeito da amostragem sistemática



$$s_L^2 = \frac{1}{2} \sum_{i=1}^n (y_i - y_{i-1})^2 / (n-1)$$

usa essa variância corrigida para  
estimar o erro padrão

# VÁRIOS OUTROS ESQUEMAS DE AMOSTRAGEM

**AMOSTRAGEM ANINHADA (Nested/Cluster)**

**AMOSTRAGEM MULTI-ESTÁGIOS**

**AMOSTRAGEM REPETIDA**

**AMOSTRAGEM ADAPTATIVA**

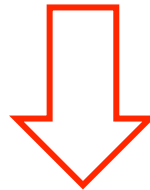
**CAPTURA-RECAPTURA**

**AMOSTRAGEM ROBUSTA -> DETECTABILIDADE**

**TAMBÉM COMBINAÇÕES ENTRE DIFERENTES ESQUEMAS**

## PRINCIPAIS REQUISITOS DE UMA BOA AMOSTRAGEM:

- UNIDADES AMOSTRAIS ALOCADAS DE FORMA ALEATÓRIA → INDEPENDENTES
- NÚMERO DE UNIDADES AMOSTRAIS DEVE SER "SUFICIENTE" (Lei dos Grandes Números)
- SEM FATORES DE CONFUSÃO



PODEMOS UTILIZAR A AMOSTRA PARA ESTIMAR OS PARÂMETROS POPULACIONAIS

**Número adequado de amostras e aleatorização devem sempre andar juntos!!**

A partir dos requisitos básicos para uma boa amostragem, vocês identificam potenciais **problemas** na amostragem anterior que realizaram? Quais?



Gostariam de realocar as 10 unidades amostrais? Indique na figura2 com o símbolo "#"

## **Atividade D)**

Expliquem, no verso da folha, que motivos levaram vocês a usar essa nova distribuição para suas 10 unidades amostrais "#"

Gostariam de realocar as 60 unidades amostrais? Indique na figura 2 com o símbolo "+"

Oeste (W)

Sul (S)



Norte (N)

Oceano Atlântico - Leste (E)

## **Atividade E)**

Expliquem, no verso da folha, que motivos levaram vocês a usar essa nova distribuição para suas 60 unidades amostrais "+"



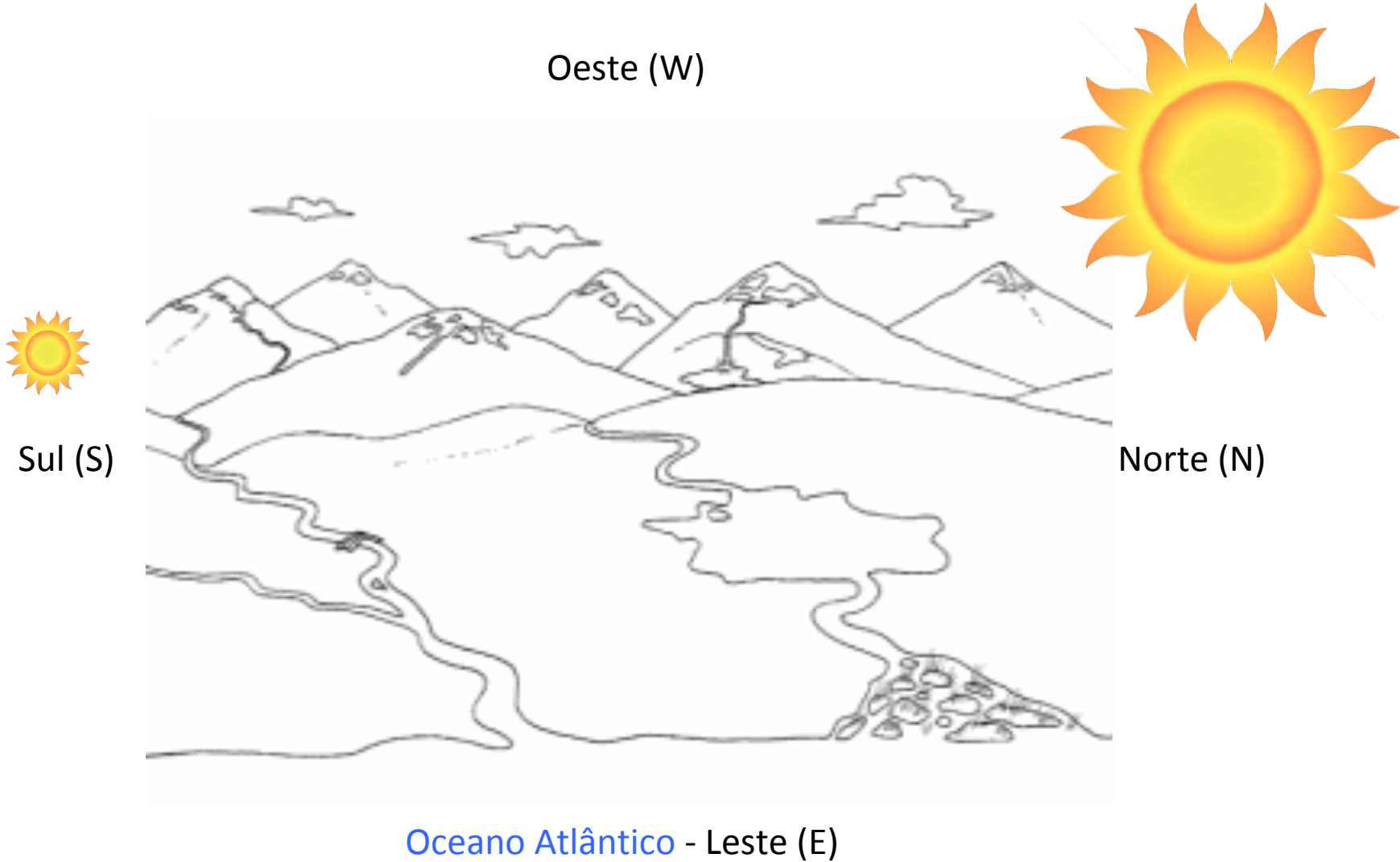
**OK, os dados foram coletados em campo dessa forma!**



Após a análise dos dados, notamos que existe uma **variação muito grande** na produção de sementes entre as unidades amostrais, mesmo considerando as 60 unidades amostrais estabelecidas.

**Quais fatores poderiam estar determinando essa variação?**

Será que pode haver diferenças entre as faces N e S das montanhas, em função do excesso de radiação incidente sobre as folhas das plantas na face N, que poderia reduzir a fotossíntese e a reprodução das plantas?



Como vocês alocariam agora as 10 unidades amostrais? Indiquem na figura 3 com o símbolo "Z"

Como vocês alocariam agora as 60 unidades amostrais? Indiquem na figura 3 com o símbolo "\*"

### **Atividade F)**

Descrevam, no verso da folha, por que essa nova distribuição das 10 unidades amostrais "z" será mais eficiente para responder a nova pergunta proposta

### **Atividade G)**

Descrevam, no verso da folha, por que essa nova distribuição das 60 unidades amostrais "\*" será mais eficiente para responder a nova pergunta proposta



**Encontramos diferenças significativas na produção de sementes entre unidades amostrais localizadas na face norte (menor produção de sementes) e na face sul (maior produção de sementes)**

**A hipótese de excesso de incidência de radiação sobre as folhas, afetando a fotossíntese e a reprodução, foi corroborada?**

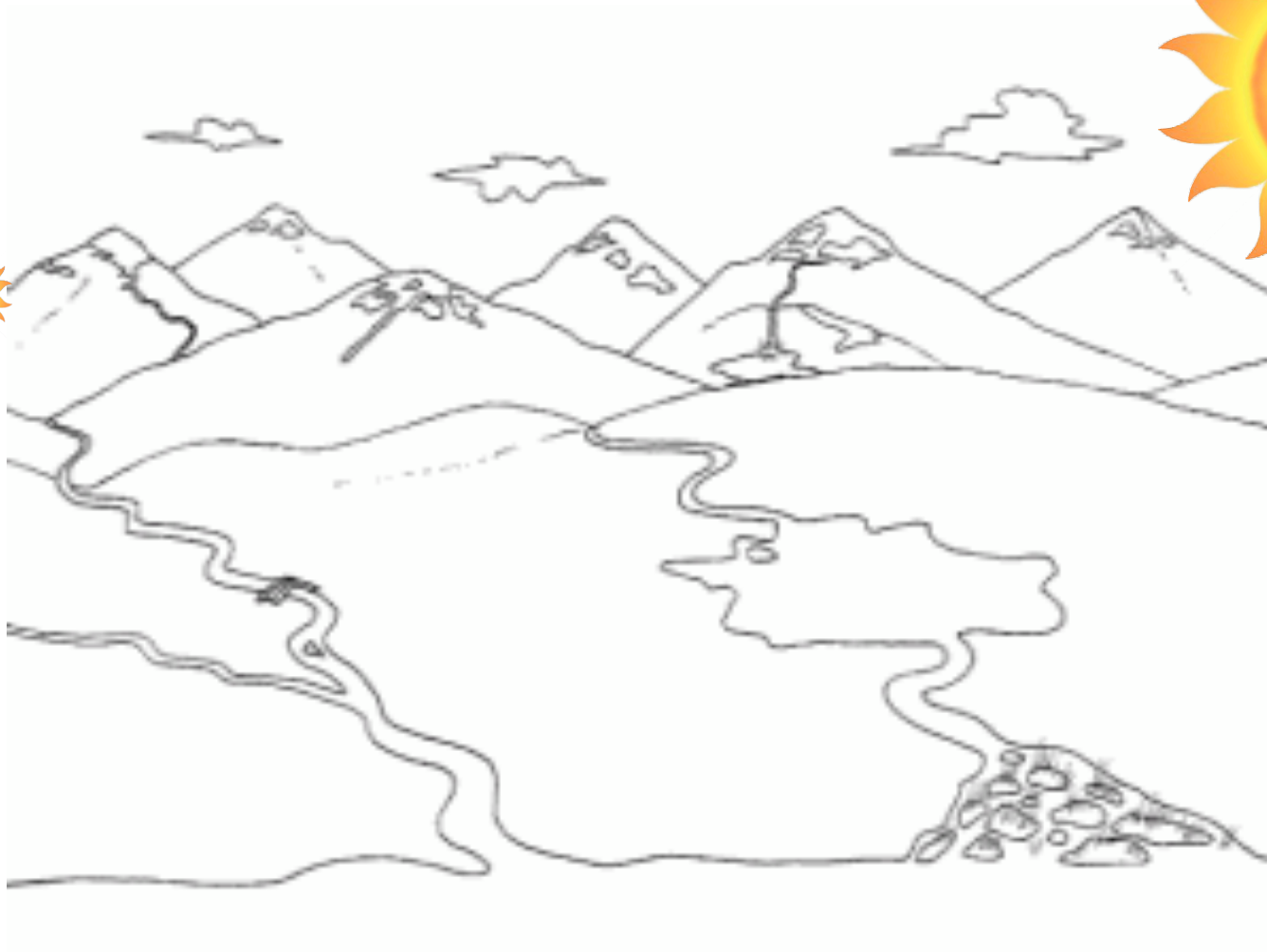
**Qual outro processo/mecanismo poderia determinar essas diferenças?**

**Será que diferenças entre faces N e S ocorrem em função do ressecamento do solo?**

Oeste (W)



Sul (S)



Norte (N)

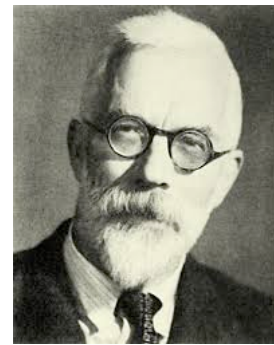
Oceano Atlântico - Leste (E)

**Como separar o efeito da radiação sobre as folhas e do ressecamento do solo?**

**DELINEAMENTOS PARA ESTUDOS EXPERIMENTAIS  
(não exclusivamente)**

# DELINEAMENTOS EXPERIMENTAIS

## O QUE MUDA?



VARIÁVEIS PREDITORAS SÃO MANIPULADAS E ISOLADAS DE OUTROS FATORES



IDEALMENTE NÃO EXISTEM FATORES DE CONFUSÃO EM UM EXPERIMENTO  
(*True-experiment*)

# TRÍADE DE PROCEDIMENTOS EXPERIMENTAIS

REPLICAÇÃO - ALEATORIZAÇÃO - **CONTROLE**

**Replicação e Aleatorização devem sempre andar juntas!!**

**Se não houver controle, não há como inferir causalidade.**

**Podem ser experimentos, mas com inferência fraca de causalidade.**

**(Holland, 1986).**



# DIFERENTES TIPOS DE CONTROLE:

## CONTROLE DE EFEITO

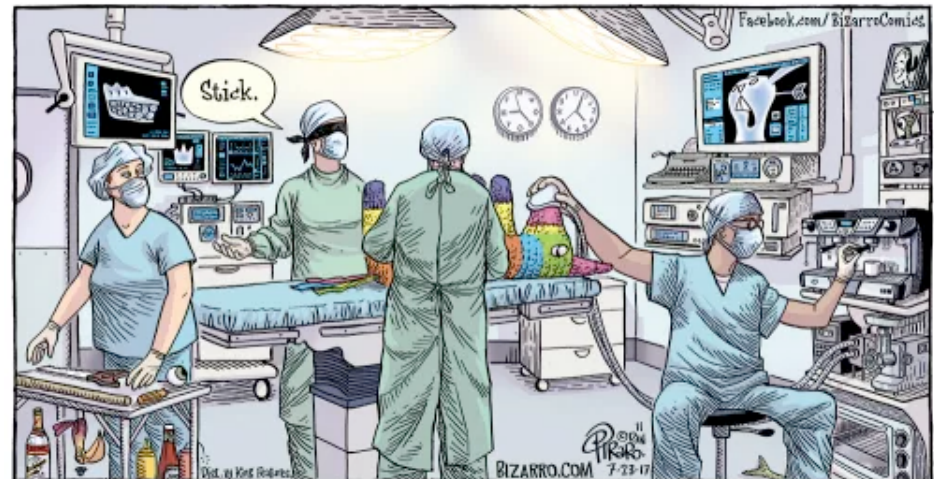
- Sem a aplicação do tratamento

## CONTROLE DE PROCEDIMENTO

- Gaiolas sem telas
- Controle osmótico (PEG)
- Garrafas sem furos
- Sham surgery



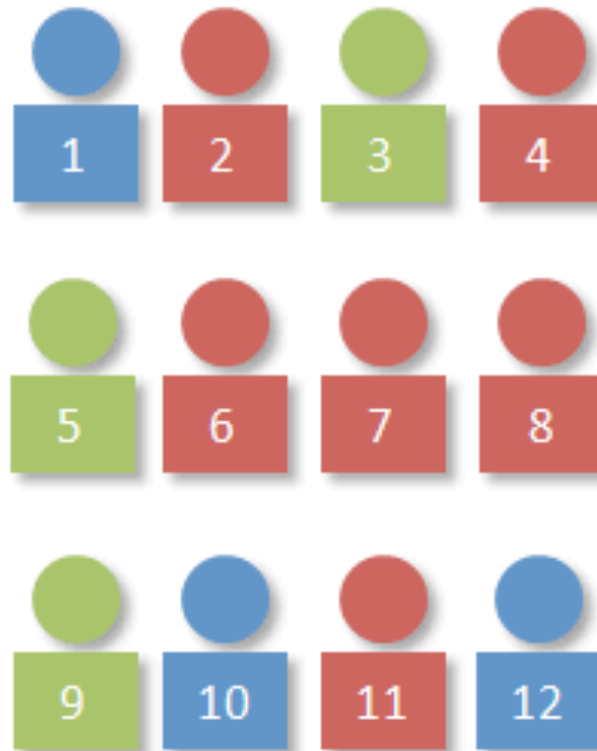
*Delphinus nuttallianum*



# ALEATORIZAÇÃO SIMPLES - UM FATOR

## PRINCÍPIO BÁSICO:

Qualquer réplica deve ter a mesma chance de ter um nível do tratamento atribuído a ela



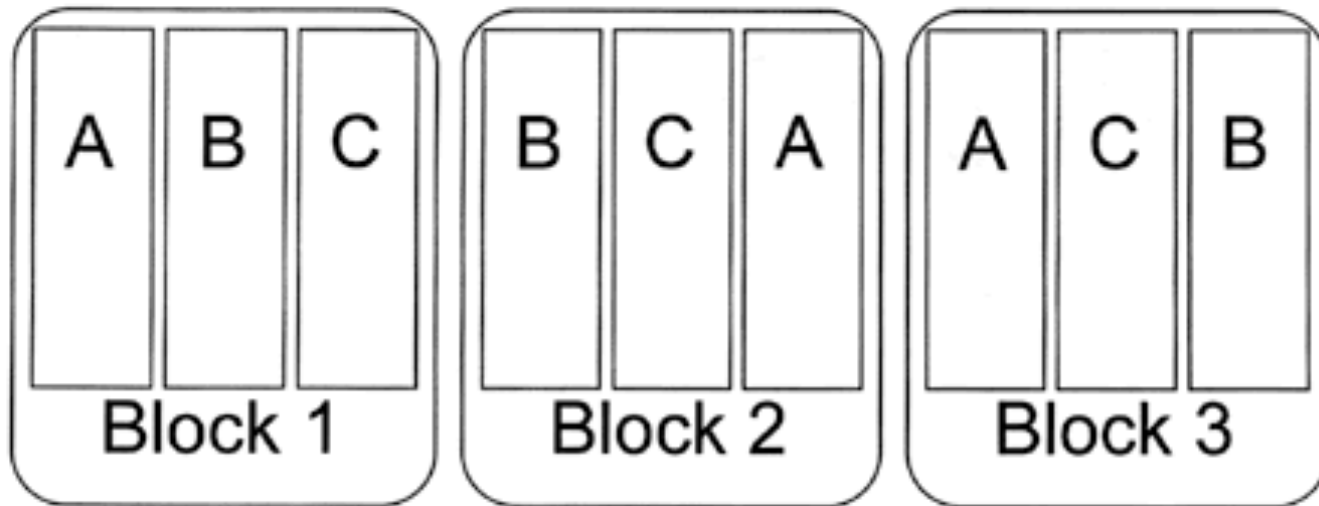
ANALITICAMENTE FÁCIL LIDAR COM NÚMEROS DIFERENTES DE RÉPLICAS POR TRATAMENTO

É o delineamento básico para a ANOVA de um fator

# BLOCOS ALEATORIZADOS

## PRINCÍPIOS BÁSICOS:

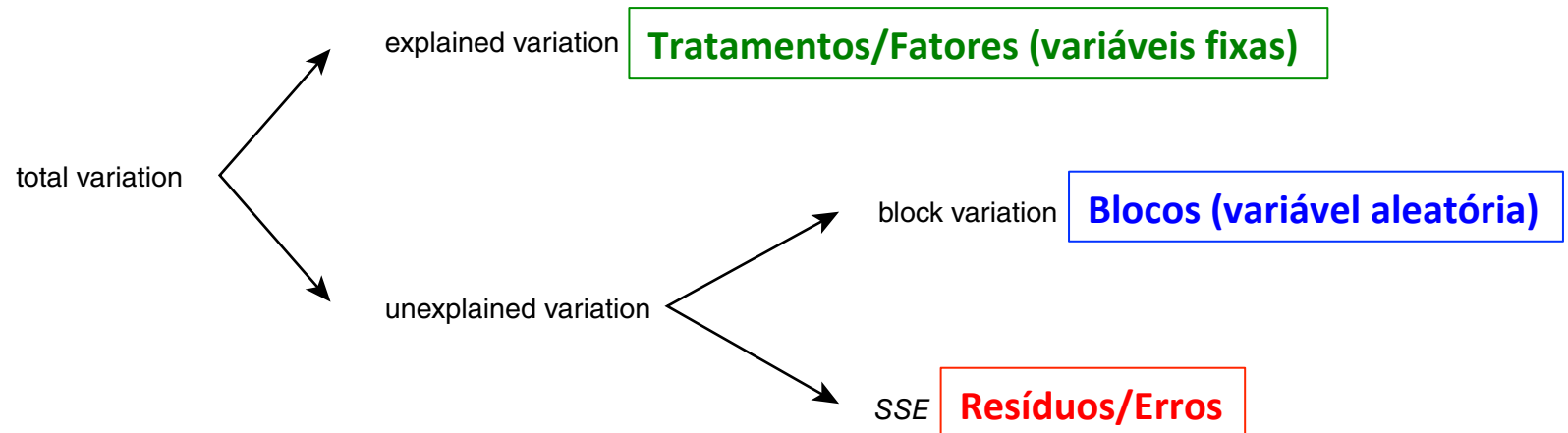
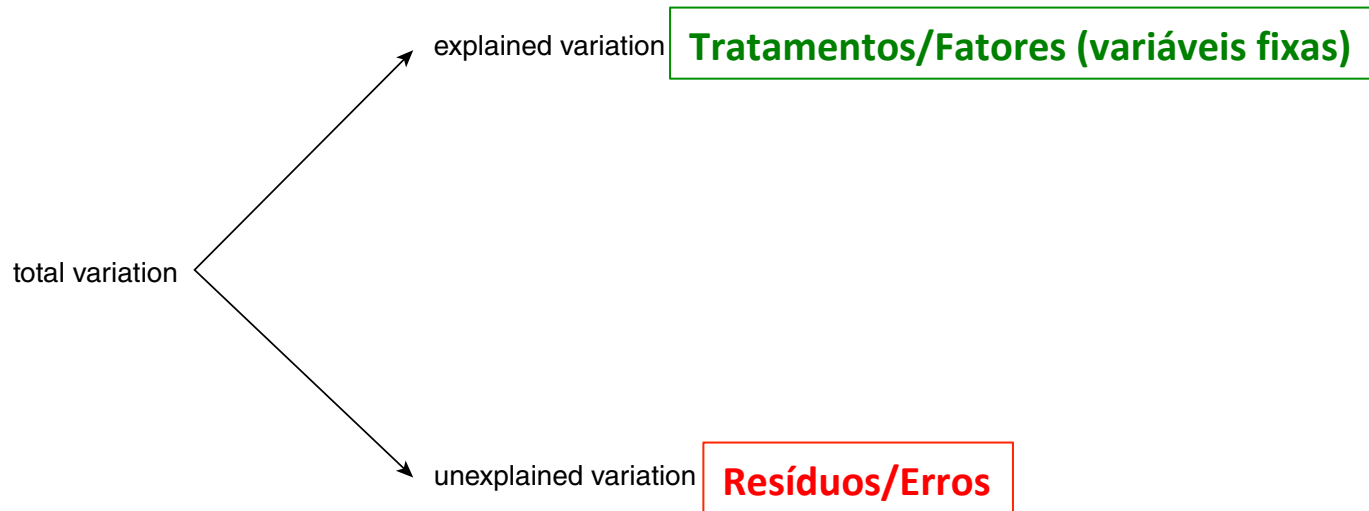
**Blocos são internamente mais homogêneos do que entre blocos**  
**Alocar tratamentos aleatoriamente dentro de blocos**



Blocos não representam um outro fator conhecido

Minimizam a chance de algum efeito desconhecido atuar em apenas um tratamento

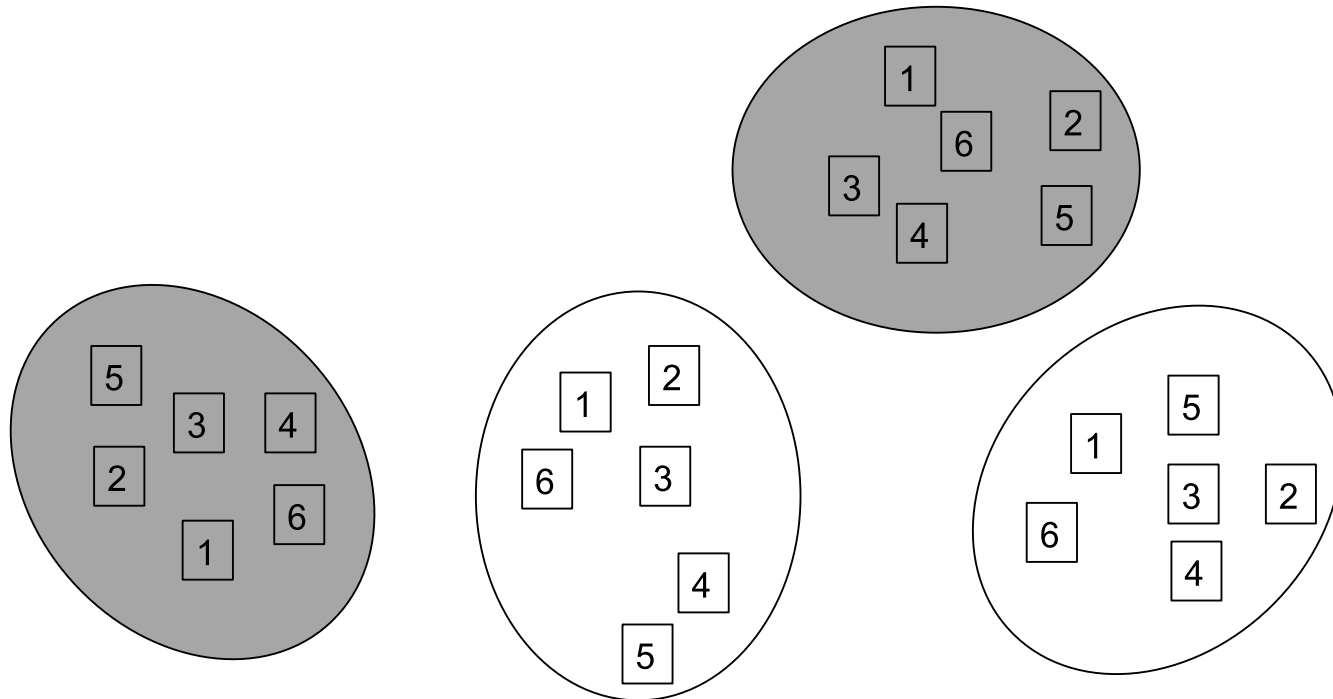
# Atribuir parte da variação aos blocos diminui a fração de variação não explicada



# SPLIT-PLOT

## PRINCÍPIO BÁSICO:

Uma réplica de cada nível de tratamento de um dos fatores do experimento inserido em cada réplica do segundo fator. Caso particular de blocos.



### Exemplo

**Fator 1:** 05 antifúngicos diferentes injetados em peixes em gaiolas e um controle (06 níveis)

**Fator 2:** 02 tipos de lagos (cinza = lagos profundos; branco = lagos rasos)

**Tamanhos amostrais ficam diferentes.**

**No exemplo acima, cada lago tem 6 réplicas, mas cada antifúngico tem apenas 4 réplicas.**

**Existe uma análise específica que lida com essas diferenças**

# MULTIFATORIAL - DOIS FATORES (com vários níveis cada)

## PRINCÍPIO BÁSICO:

Os tratamentos são completamente cruzados e ortogonais. Todo nível de tratamento do primeiro fator é representado com todos os níveis do segundo

**Two way ANOVA**  
**2 factors: Row vs. Column factors (R X C)**

COLUMNS:      C levels

	1	2	...	C
1	<i>Group11</i>	<i>Group12</i>	...	<i>Group1C</i>
2	<i>Group21</i>	...	...	...
...	...	...	...	...
R	<i>GroupR1</i>	...	...	<i>GroupRC</i>

ROWS: R levels

**Cuidado com o ESFORÇO AMOSTRAL ao incluir fatores e/ou níveis dentro dos fatores.  
Aumentos muito rápidos (exponencial)!**

# MUITOS OUTROS DELINEAMENTOS

DESIGN TYPE	SCHEMA
A-1 Completely Randomized	
A-2 Randomized Block	
A-3 Systematic	
B-1 Simple Segregation	
B-2 Clumped Segregation	
B-3 Isolative Segregation	
B-4 Randomized, but with inter-dependent replicates	
B-5 No replication	

Hurlbert (1984)

**E MAIS:  
SÉRIES TEMPORAIS  
ANTES-DEPOIS-CONTROLE-IMPACTO (ADCI)**

...

## Porque um bom delineamento ajuda?

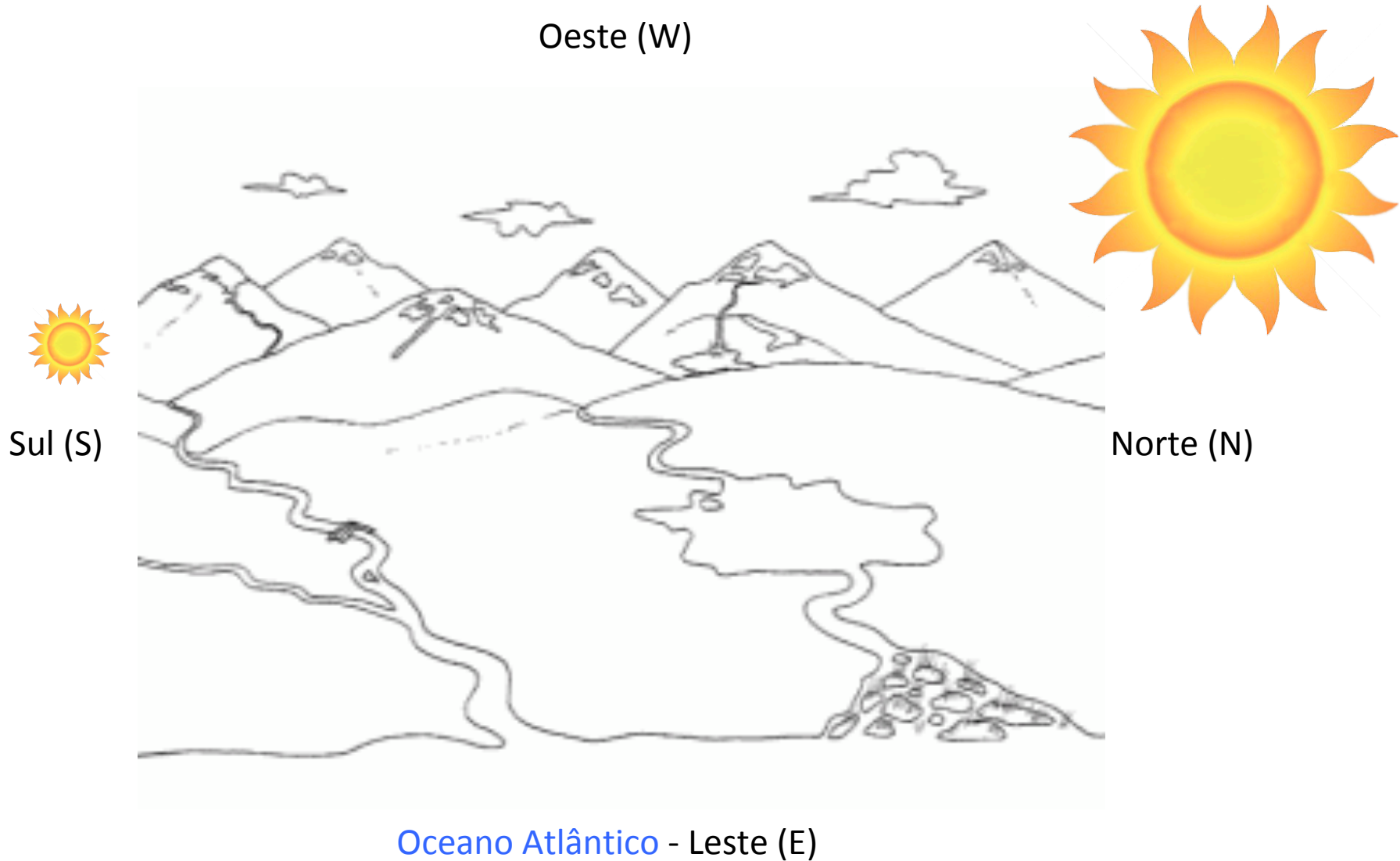
- Melhor distribuição do esforço amostral
- Aumento do poder e robustez do teste estatístico
- Identificação clara de quais tratamentos (e níveis) são essenciais
- Aumento do potencial de generalização dos resultados
- Elimina ou minimiza fatores de confusão
- Diminui a variação não explicada
- **Minimiza as chances de erros Tipo I e Tipo II**



## Resumo e algumas dicas importantes:

- A definição de o que é um bom delineamento depende da pergunta do estudo;
- Sempre será fundamental ter a população estatística definida *a priori* e claramente;
- Muitas decisões são baseadas em experiência;
- Se não tiver experiência, procure a literatura;
- Todas as decisões devem ser embasadas em algo concreto, que possa ser justificado;
- Prefira sempre a aleatorização do que a conveniência na hora de definir unidades amostrais/experimentais(Incorpore a aleatorização);
- Estratificar é bom, mas o delineamento precisa ser incorporado na análise adequada;
- Se os planos iniciais não derem certo no campo/laboratório pare e reflita novamente sobre as implicações de cada decisão;
- Quando estiver em campo/laboratório e algo não funcionou como planejado durante a coleta de dados, sempre anote as mudanças realizadas;
- Um bom delineamento consegue até prever se uma moeda vai dar cara ou coroa (ver artigo sobre o Euro e dos Bayesianos).

O menor número de sementes produzido pelas plantas localizadas na face norte é explicado pelo efeito do excesso de radiação sobre as folhas ou pelo efeito do ressecamento do solo?



## **Atividade H)**

Descrevam, no verso da folha, um experimento para separar o efeito da incidência solar sobre as folhas ou sobre o solo





# MESMO COM EXPERIMENTOS, É DIFÍCIL DETERMINAR O MECANISMO DEFINITIVO

## Decisão é filosófica:

Onde parar?

Qual é o limite?

Qual mecanismo é suficiente para a aplicação dos resultados (custos x benefícios)?

### CASO DAS BACTÉRIAS E TOXINAS

- Experimento muito bem delineado demonstra relação entre presença da bactéria A e a doença "mal do estatístico". Nos resultados existe uma certa variação, mas dentro dos limites aceitos.
- Bactéria A tem a toxina X, mas nem sempre; Doentes nem sempre têm bactérias A e são podem ter;
- Bactéria B também tem toxina X, mas nem sempre (variação nos resultados de manifestação da doença). Resultados similares aos da bactéria A

**Mecanismo: Bactérias ou toxinas? Faz diferença na aplicação? Na Profilaxia?**

## Outros exemplos

- Alelopatia sem controlar PEG
- Adição de fertilizantes sem analisar alteração de pH do solo (i.e. não foi a fertilidade e sim o pH que promoveu mudanças)

## Outros problemas

### Exemplo de manipulação excessiva:

*Clones seriam bons para fazer experimentos de efeitos de remédios?*

Talvez não, pois não exprimem a variação entre indivíduos.

Mas servem para quantificar a variabilidade devido a outros fatores (ambientais, manipulação, etc)