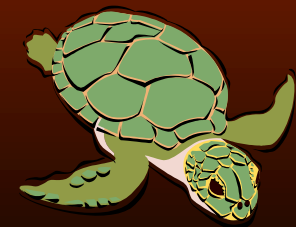
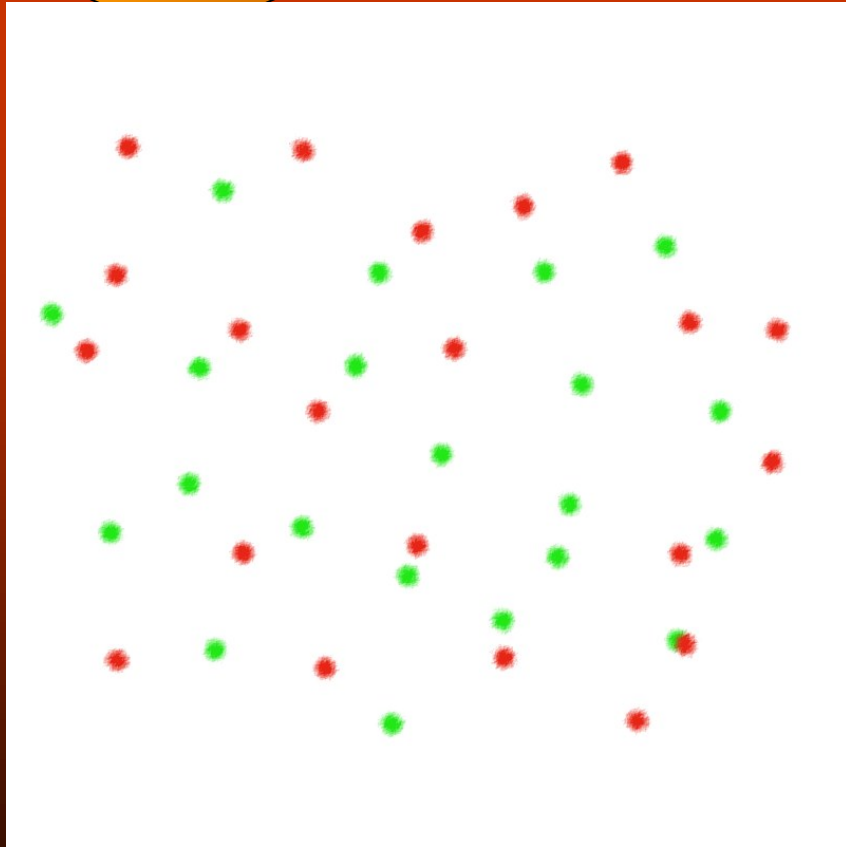


*Métodos de Amostragem de
Populações de Animais Silvestres*

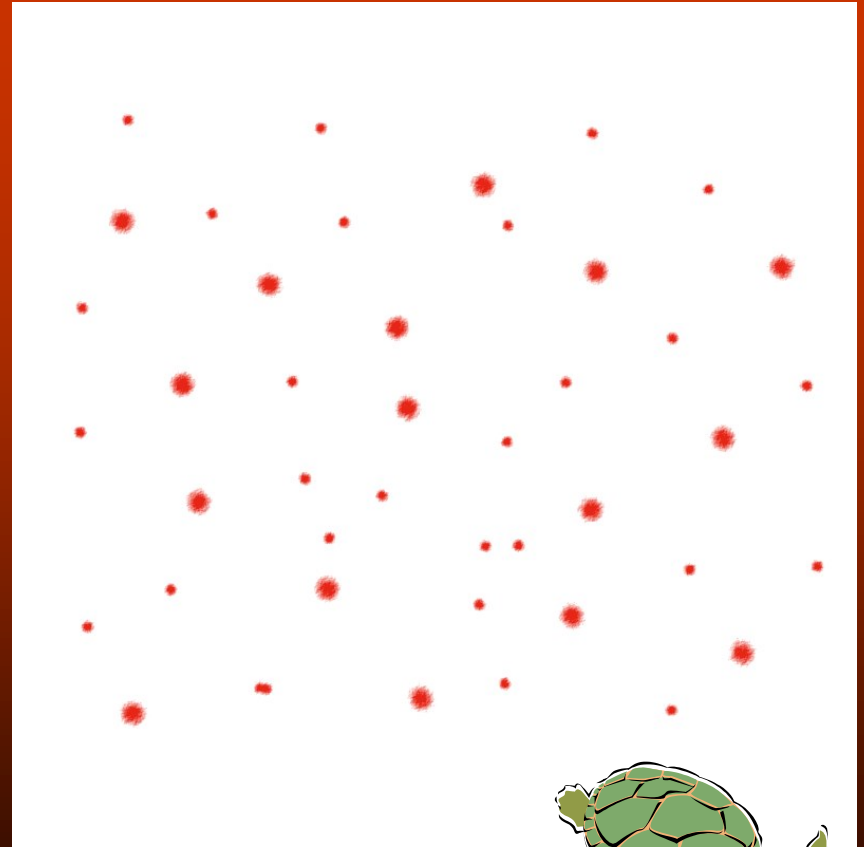
*Amostragem de
variável binomial*



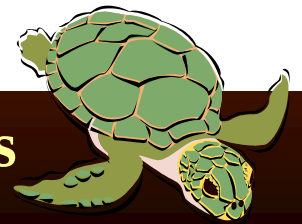
O que é uma variável binomial?



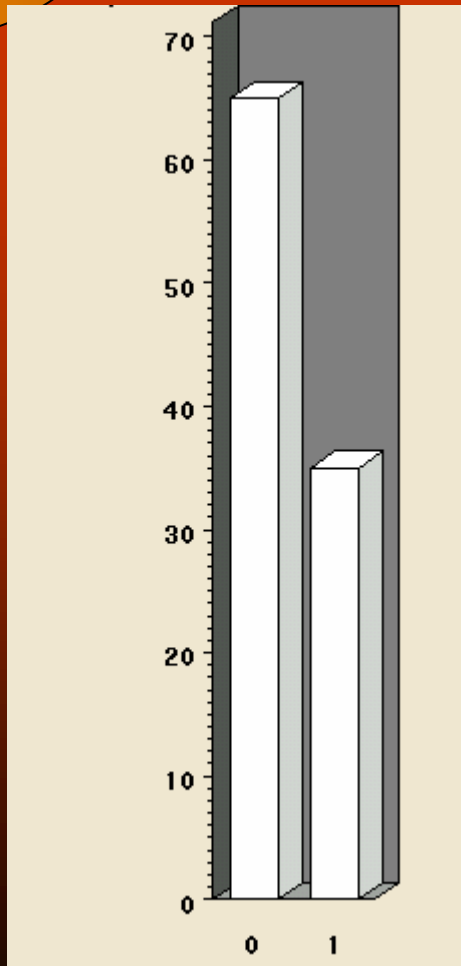
2 qualidades



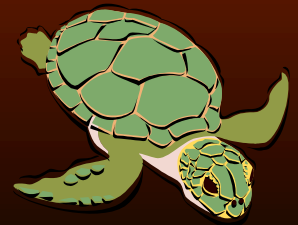
2 tamanhos



Distribuição de freqüência:



0	0	0	0	0
0	1	1	0	0
0	0	0	0	0
0	0	1	0	0
1	0	0	0	0
1	1	1	0	0
1	0	1	0	0
0	0	1	0	1
0	0	1	0	1
1	0	0	0	1



Média:

0

0

0

0

1

1

1

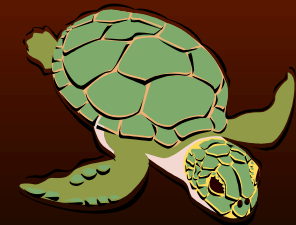
0

0

1

$$\bar{x} = \frac{\sum x_i}{n} = \frac{0 + 0 + 0 + 0 + 1 + 1 + 1 + 0 + 0 + 1}{10} = \frac{4}{10} = 0,4$$

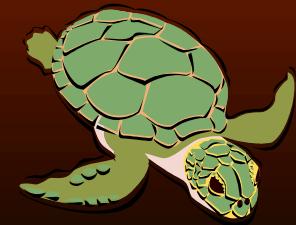
$$\hat{p} = \bar{x} = 0,4 = \textit{proporção}$$



Variância da média:

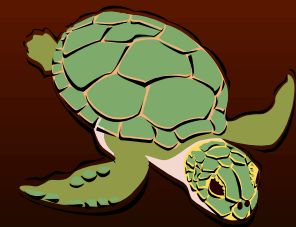
$$s_{\hat{p}}^2 = \frac{\cancel{n}\hat{p}(1 - \hat{p})}{\cancel{n}(n - 1)} \cdot \left(\frac{N - n}{N}\right)$$

$$s_{\hat{p}}^2 = \frac{\hat{p}(1 - \hat{p})}{(n - 1)} \cdot \left(\frac{N - n}{N}\right)$$



Intervalo de confiança:

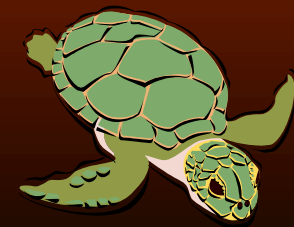
$$\hat{p} \pm t \cdot \sqrt{s_{\hat{p}}^2}$$



Erro da amostragem:

$$EA\% = \frac{100.t.\sqrt{s_{\hat{p}}^2}}{\hat{p}}$$

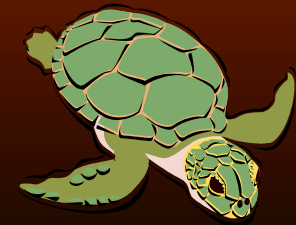
t com (n-1) g.l.



Intensidade da amostragem:

$$n^* = \frac{t^2 \cdot N \cdot \hat{p} \cdot (1 - \hat{p})}{d^2 \cdot N + t^2 \cdot \hat{p} \cdot (1 - \hat{p})}$$

d = erro desejado (valor absoluto)



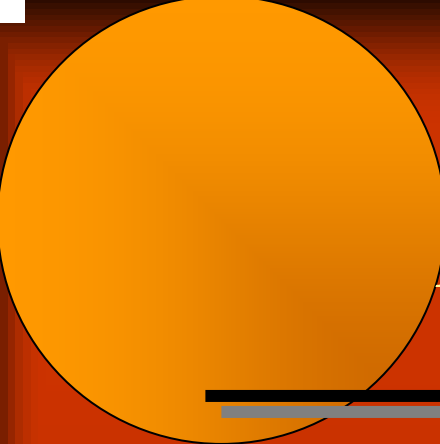
Exercício:

Você deseja conhecer o número de ninhos de pica-pau em árvores que apresentam oco. Para tanto fez-se um levantamento em uma área de 100 ha.

Realizou-se uma amostragem de 200 árvores ao acaso e entre elas 35 possuíam ninhos ativos. Um levantamento do habitat mostrou que a área possui em média 689 árvores por hectare. Fazer um programa SAS para calcular os seguintes valores:

- a) Número total de ninhos na área.
- b) Erro da amostragem
- c) Intensidade da amostragem para um erro de 0,05 e probabilidade de 95%

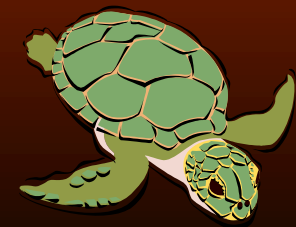




Diversas proporções:

Estimativa simultânea de diversas
proporções

Thompson, Sampling, (1992)



Exemplo:

Uma bióloga necessita coletar uma amostra de peixes para estimar a proporção da população em cada classe de idade e ela deseja que o intervalo de confiança tenha 95% de probabilidade e seja de 0,05 das proporções estimadas. Ela não sabe quantas classes de idade existem na população. Qual o tamanho da amostra aleatória que ela deve selecionar?

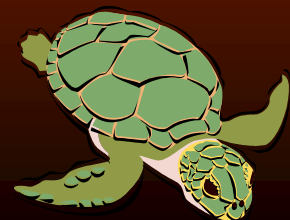


Tabela:

α	$d^2 n^*$	n^* para $d=0,05$	m
0,50	0,44129	177	4
0,40	0,50729	203	4
0,30	0,60123	241	3
0,20	0,74739	299	3
0,10	1,00635	403	3
0,05	1,27359	510	3
0,025	1,55963	624	2
0,02	1,65872	664	2
0,01	1,96986	788	2
0,005	2,28514	915	2
0,001	3,02892	1212	2
0,0005	3,33530	1342	2
0,0001	4,11209	1645	2

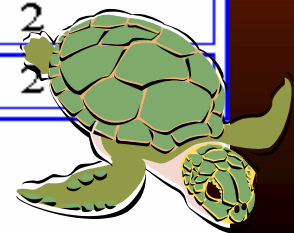


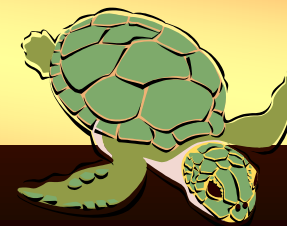
Tabela:

Nota: α = probabilidade da estimativa estar fora do intervalo de confiança ($1 - \alpha$ = probabilidade da estimativa estar dentro do intervalo de confiança).

d = valor do intervalo de confiança desejado (0,05, p. ex.).

n^ = tamanho da amostra.*

m = número máximo de empates e as demais proporções sendo zero.



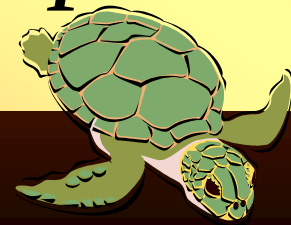
Cálculos:

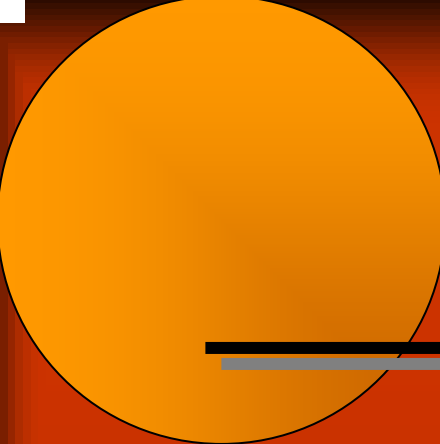
$$d = 0,05$$

$$\alpha = 0,05$$

$$d^2 n^* = 1,27359$$

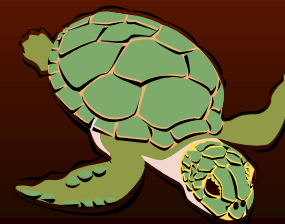
$$n^* = \frac{1,27359}{0,05^2} = 510 \text{ peixes}$$

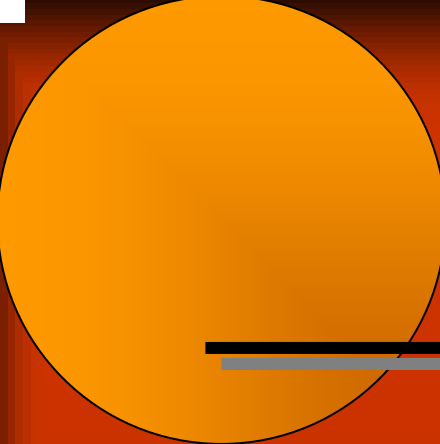




Se ela conhecesse N :

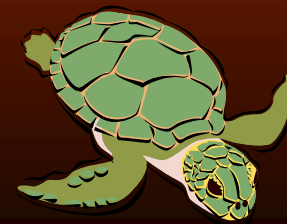
$$n^{**} = \frac{1}{\frac{1}{n^*} + \frac{1}{N}} = \frac{1}{\frac{1}{510} + \frac{1}{2000}} = 406 \text{ peixes}$$





Exercício:

Fazer um programa em SAS para calcular os dados do exemplo anterior. Colocar a tabela no programa de modo que se torne mais geral e você não precise toda vez consultar a tabela.



Amostragem sistemática:

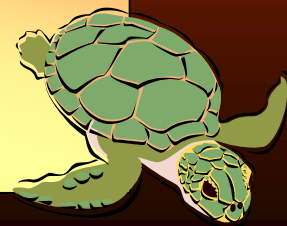
A amostragem sistemática envolve a seleção aleatória de um elemento entre os k primeiros elementos e então a seleção é feita a cada k elementos. k é chamado de intervalo de amostragem e é calculado como:

$$k = \text{Int}\left(\frac{N}{n}\right),$$

onde : Int = inteiro

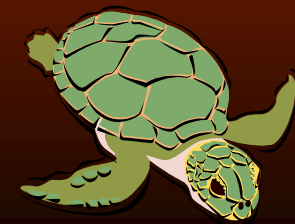
$N = \text{elementos_na_população_}(\text{tamanho_da_população})$

$n = \text{elementos_na_amostra}(\text{tamanho_da_amostra})$



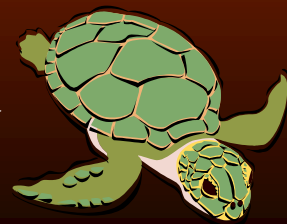
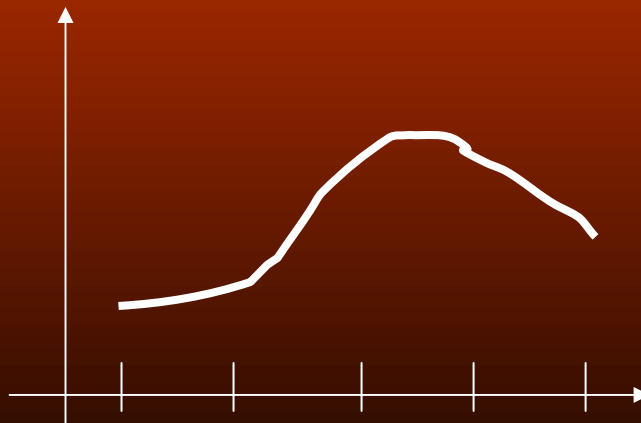
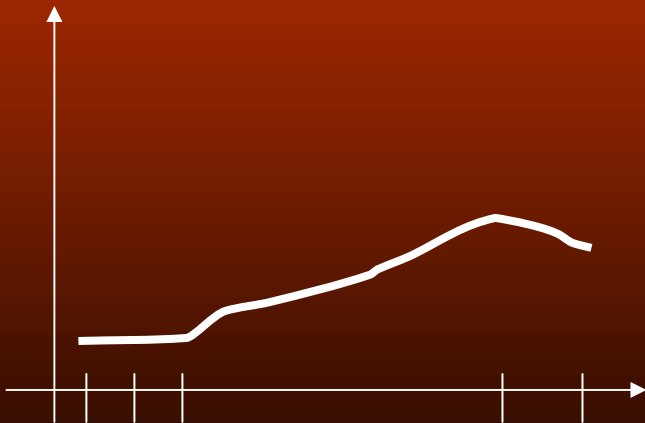
Razões para uso:

1. Em muitas situações a amostragem sistemática é muito mais efetiva em relação ao custo de coleta de dados do que a ASA.



Razões para uso:

2. Dá mais informação para cada amostra, pois os elementos de cada amostra são distribuídos de forma mais homogênea na área.

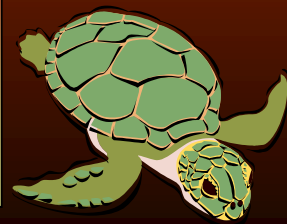


Razões para uso:

3. Pode ser usada quando não se conhece a estrutura da amostragem antes do levantamento.

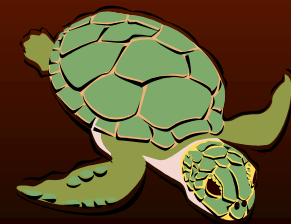
Ex.: Estamos interessados em conhecer o dano causado por uma determinada praga em uma plantação de laranja. Não conhecemos o número de pés de laranja na área de estudo. Avaliamos 1 pé a cada 500 e no final temos 35 pés avaliados. Quantos pés de laranja temos na área de estudo?

$$k = \frac{N}{n} \implies N = n.k = 35 \times 500 = 17.500 \text{ pés}$$

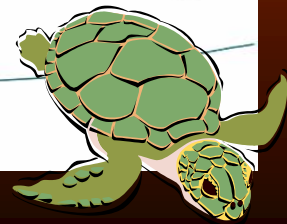
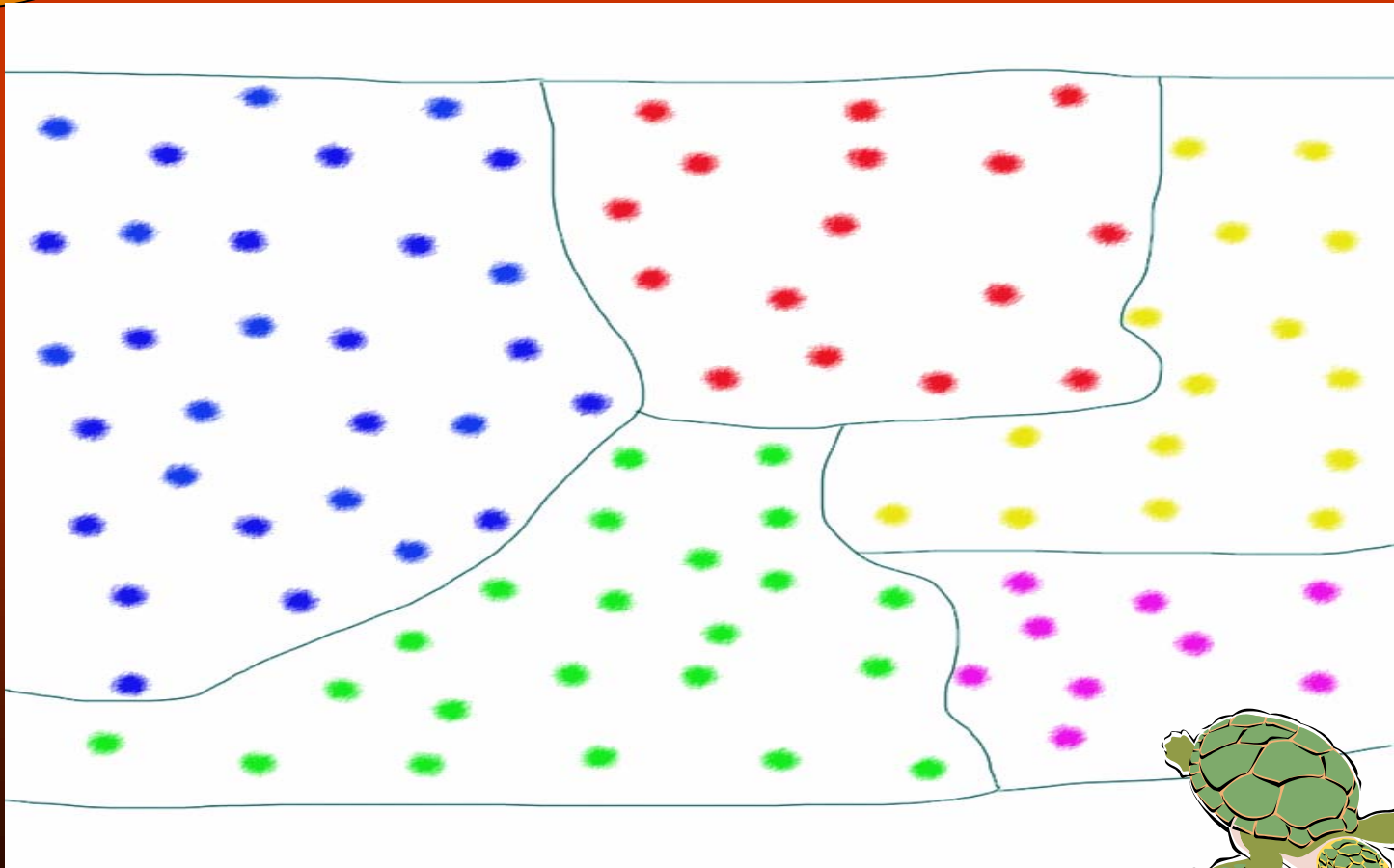


Razões para uso:

4. As mesmas fórmulas usadas para ASA podem ser usadas para AS, para estimar a média da população, o total e a proporção.
5. Possibilita o uso dos dados da amostragem (principalmente quando geo-referenciados – GPS) para a elaboração de mapas de distribuição espacial.

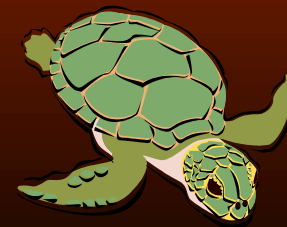


Mapa de distribuição espacial

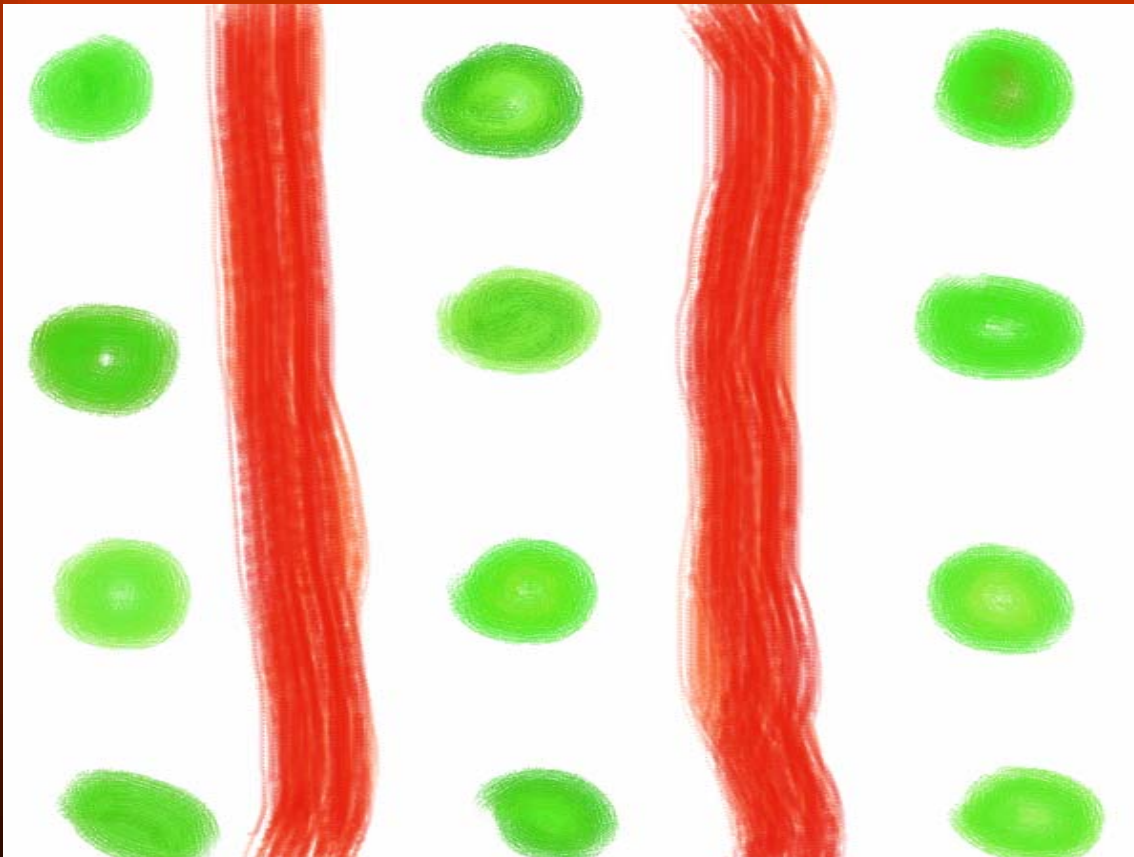


Desvantagens:

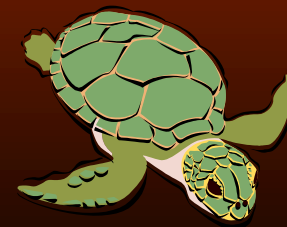
1. A estimativa da variância é enviesada (sempre maior).
2. Quando a população é periódica as estimativas da média e variância são também enviesadas.



Desvantagens:



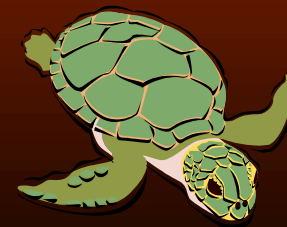
*População
periódica*



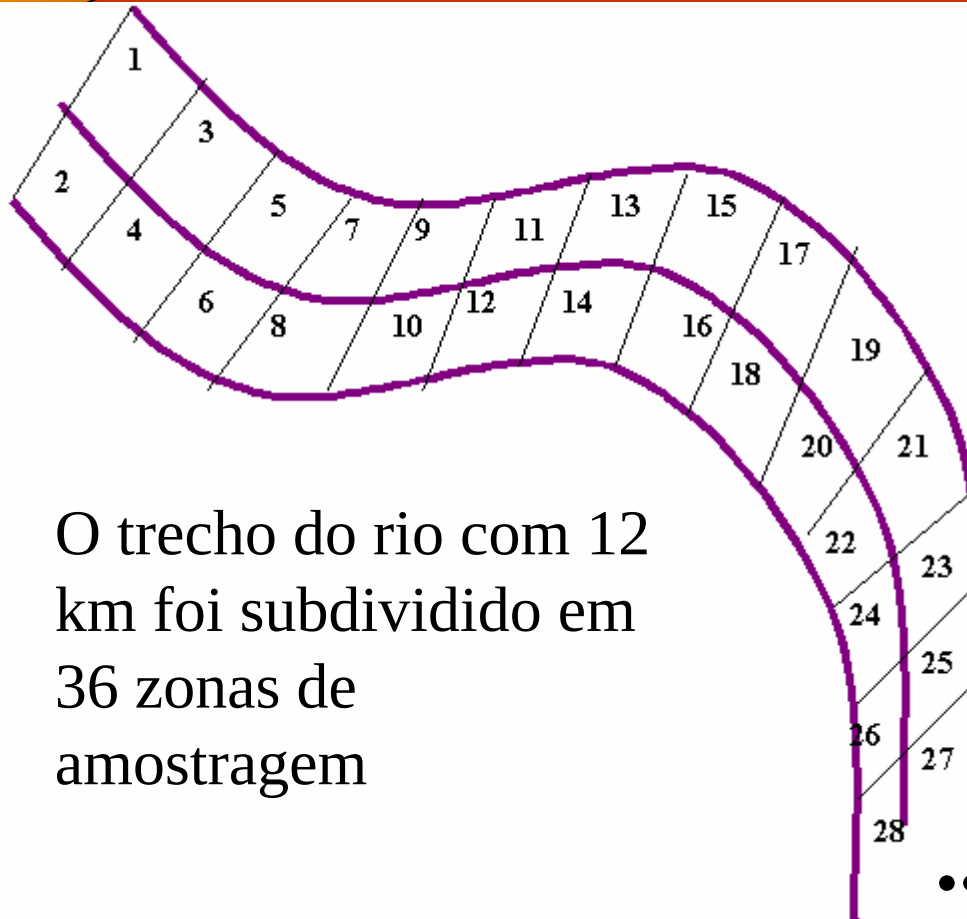
Exercício:

Um estudo é planejado para avaliar o nível do pesticida dieldrin, que é considerado um suposto agente cancerígeno, num trecho de 12 km de um rio. Para assegurar a representatividade o referido trecho é dividido em 36 zonas e a amostragem sistemática da água para análise foi realizada.

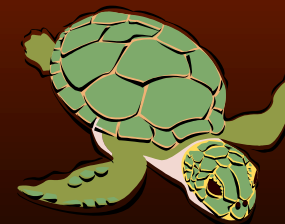
Os valores do teor de dieldrin são apresentados em micrograma por litro.



Exercício:



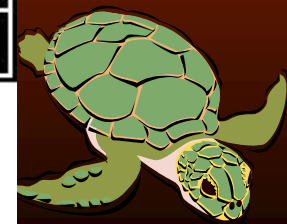
O trecho do rio com 12 km foi subdividido em 36 zonas de amostragem



Exercício:

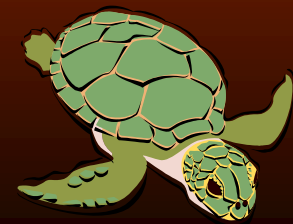
Trecho	Dioldrin	Trecho	Dioldrin	Trecho	Dioldrin
1	0	13	5	25	2
2	0	14	6	26	2
3	0	15	5	27	2
4	1	16	6	28	1
5	1	17	5	29	1
6	0	18	5	30	1
7	2	19	3	31	1
8	1	20	4	32	1
9	2	21	2	33	1
10	2	22	4	34	1
11	5	23	2	35	0
12	6	24	2	36	0

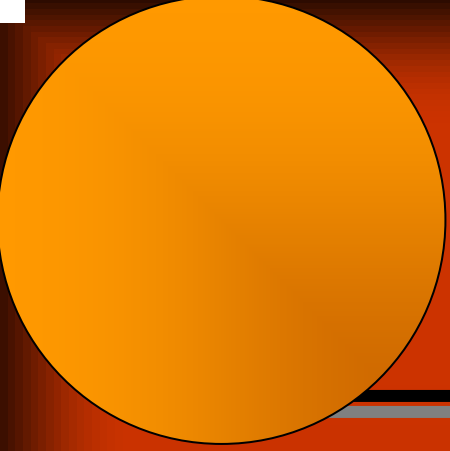
Trecho = zona do rio



Exercício:

Fazer um programa SAS para selecionar uma amostra sistemática a cada 4 zonas (trecho) para estimar a média, o intervalo de confiança, o erro e a intensidade da amostragem para um erro desejado de 10% com 95 % de probabilidade.





**Obrigado e
até a próxima
aula !!!**

