

O problema da amostragem em ecologia vegetal

Valério De Patta Pillar
Departamento de Botânica, UFRGS

A adequação dos procedimentos de amostragem depende dos objetivos do trabalho e das características do meio a ser amostrado. Em ecologia vegetal o meio nem sempre se enquadra nas condições ideais tratadas pela estatística convencional. Este texto define alguns termos e busca distinguir diferentes objetivos e respectivas soluções para a amostragem em ecologia vegetal.

Introdução

A amostragem é necessária quando não é possível ou não é conveniente acessar a totalidade de um universo amostral. Assim, tomam-se informações sobre uma parte deste, uma *amostra*, para inferir atributos sobre o seu todo. Um *universo amostral* ou *população* é aqui definido como um conjunto cujos membros são do mesmo tipo. As unidades que compõem o universo amostral, ou seja, as *unidades amostrais*, podem ser objetos perfeitamente distinguíveis, como uma planta isolada, uma ilha, um lago, etc., ou agregados de objetos, como uma área de vegetação, um perfil de solo, um volume de água, etc. Em ecologia de vegetação, usa-se o termo "quadro" ou "unidade de comunidade vegetal" para designar uma unidade amostral. Uma amostra, portanto, é composta por um certo número de unidades amostrais, geralmente muito inferior ao número total de unidades amostrais no universo amostral.

Ao amostrar é necessário decidir quanto ao número, seleção (aleatória, sistemática, ou preferencial) e, em alguns casos, tamanho e forma das unidades amostrais que irão compor a amostra. Essas decisões devem ser guiadas pelo contexto, especialmente em ecologia vegetal, em que o meio amostrado nem sempre se enquadra nas condições ideais tratadas pela estatística convencional. Kenkel, Juhász-Nagy & Podani (1989) reconhecem diferentes contextos de amostragem, resultado da combinação de dicotomias: (1) estimativa de atributos simples versus estudo de padrões e (2) universo de amostragem discreto versus universo de amostragem contínuo (Tabela 1). Alguns desses casos são proeminentes nos estudos ecológicos.

Estimativa de atributos simples versus estudo de padrões

O objetivo da amostragem pode ser, por um lado, estimar o estado de um ou mais atributos simples do universo amostral através da observação de várias unidades amostrais. Estimativas assim obtidas podem ser usadas para comparar universos amostrais. Geralmente estimam-se médias e variâncias quando os atributos são quantitativos. Exemplo típico é a descrição de um quadro de vegetação quando um censo é impossível ou inconveniente; nesse caso a descrição requer uma amostragem cujo objetivo é estimar atributos simples, ou seja, estimar a composição florística expressa em biomassa, cobertura, densidade ou frequência de cada espécie. São também casos de amostragem com esse objetivo, por exemplo, estimar numa área de vegetação a densidade média de afilhos de uma gramínea ou a biomassa aérea total por unidade de área; ou estimar em plantas de uma determinada espécie o número de sementes por inflorescência; ou estimar as condições médias do solo como pH, textura, teores de nutrientes.

O objetivo da amostragem pode ser, por outro lado, confirmar ou detectar padrões estruturais e suas conexões, geralmente envolvendo a derivação de atributos complexos a partir de atributos simples avaliados num conjunto de unidades amostrais. São exemplos a classificação, mapeamento e ordenação de comunidades vegetais, a correlação entre fatores de solo e produtividade, a correlação entre luminosidade e cobertura com uma dada espécie, ou entre vegetação (composição florística) e um conjunto de fatores de ambiente.

Universo de amostragem discreto versus universo de amostragem contínuo

Quando o universo de amostragem é discreto as unidades amostrais são naturais, distintas, reconhecíveis, tais como plantas individuais ou unidades geográficas isoladas (ilhas, lagos). O universo de amostragem assim definido tem um tamanho finito e um número finito de amostras possíveis e se pode fazer uma lista das unidades amostrais. A amostragem nesse caso é apenas uma questão de definir o número e o método de seleção das unidades amostrais.

Quando o universo de amostragem é contínuo as unidades amostrais são agregados arbitrariamente delimitados, como uma área de vegetação, um perfil de solo, uma tradada de solo, um volume de água, etc, havendo teoricamente um número infinito de unidades amostrais possíveis. Note-se que nesse caso a amostragem também envolve definir os limites desses agregados (área, volume, forma), e essas decisões podem influir nos resultados da amostragem.

Tabela 1. Exemplos de contextos de amostragem (adaptado de Kenkel, Juhász-Nagy & Podani, 1989).

	Objetivo	Universo de amostragem	Exemplos de atributos avaliados
1	Estimativa de atributos simples	Discreto	- Altura média das árvores - Produção de sementes por planta - Médias de variáveis morfológicas, como altura de planta, largura de folha, etc.
2		Contínuo	- Cobertura ou biomassa de quadros de vegetação como um todo ou de cada espécie - Teores médios de nutrientes e pH do solo
3	Estudar padrões estruturais (atributos complexos)	Discreto	- Padrão de distribuição espacial de indivíduos de uma espécie - Estado da ordenação e classificação de entidades discretas, como OTUs
4		Contínuo	- Padrão de distribuição espacial da cobertura de uma espécie - Estado da ordenação e classificação de comunidades vegetais - Correlação entre dois conjuntos de variáveis, e.g., a correlação entre vegetação e ambiente

Tamanho e forma da unidade amostral

Sabe-se que a precisão da estimativa de uma média aumenta na medida em que se reduz a variabilidade (variância) entre as unidades amostrais ou se aumenta o número de unidades amostrais (o tamanho da amostra). Portanto, quando o universo de amostragem é discreto, a

precisão da estimativa é apenas uma questão de ajuste no número de unidades amostrais, já que a variância é intrínseca ao universo amostral. Porém, quando o universo amostral é contínuo, com um número infinito de possíveis unidades amostrais, a variância vai depender do tamanho e forma das unidades amostrais. Vejamos um exemplo: Numa dada área de vegetação o padrão de distribuição espacial da composição florística mostra zonas em que uma espécie apresenta-se com maior cobertura e outras em que a mesma espécie apresenta-se com menor cobertura; se usamos um quadro de forma quadrada ou circular e de tamanho semelhante ao dos agregados de maior cobertura, e localizamos um certo número desses quadros de inventário aleatoriamente sobre a área de vegetação, é provável que a cobertura observada em alguns desses quadros será mínima e em outros quadros será máxima, resultando numa variância relativamente alta (Greig-Smith 1957, Podani et al. 1993). O contrário se espera que aconteça se usamos um quadro de forma mais alongada e/ou de tamanho maior de forma a abranger zonas de maior e menor densidade no mesmo quadro, porém os padrões de ocorrência da espécie não serão percebidos tão nitidamente. Portanto, se o objetivo é estimar algum atributo simples, como a cobertura média da espécie na área de vegetação estudada, aumenta-se a precisão da estimativa, para um mesmo número de quadros, usando quadros maiores e/ou mais alongados, heterogêneos internamente. Mas, se o objetivo é detectar padrões, o procedimento é completamente o oposto: quadros relativamente pequenos, de forma quadrada ou circular, homogêneos internamente, são indicados. Portanto, as condições de amostragem que satisfazem o objetivo de estimar atributos simples podem não coincidir com as que satisfazem o objetivo de estudar padrões (Orlóci & Pillar 1989). A bibliografia estatística e ecológica geralmente é falha na distinção desses objetivos.

Seleção das unidades amostrais

A amostragem é *aleatória* quando todas as unidades amostrais têm a mesma probabilidade de serem incluídas na amostra. Para tanto, é necessário ter uma lista ou mapa de todas as unidades amostrais, as quais podem ser então tomadas aleatoriamente. A amostragem é *aleatória estratificada* quando o universo amostral é dividido em estratos, segmentos, o que pode ser feito de forma subjetiva, e em cada estrato é feita a seleção aleatória das unidades amostrais. A amostragem é *sistemática* quando apenas o primeiro membro da amostra é selecionado ao acaso, sendo os demais tomados a intervalos regulares. Para estimar atributos simples a seleção das unidades amostrais deve seguir um desses critérios, pois do contrário a estimativa do atributo será viciada.

No entanto, grande parte do que se sabe a respeito de processos biológicos, organismos, populações e comunidades é resultado de pesquisas em que foi usada amostragem *preferencial*, em que as unidades são selecionadas porque parecem típicas ao pesquisador (Orlóci 1991). Nesse caso, ou a propriedade considerada é uniforme na população (e.g., número de cromossomas) não sendo importante o método de seleção das unidades estudadas, ou o objetivo da amostragem é confirmar padrões mais ou menos evidentes. Por exemplo, em taxonomia usam-se espécimens tipo; em fitossociologia, e.g. Braun-Blanquet (1928), selecionam-se sítios homogêneos para delimitar a comunidade vegetal a ser descrita, porque padrões de vegetação são muitas vezes óbvios ao pesquisador, sendo mais eficiente descrever cada mancha onde parecem mais típicas,

homogêneas, do que descrever e analisar um sem número de unidades amostrais aleatórias para descrever o que é óbvio. A amostragem preferencial portanto é plenamente aceitável em ecologia quando se objetiva estudar ou confirmar padrões percebidos subjetivamente.

Para Mueller-Dombois & Ellenberg (1974) a amostragem de comunidades vegetais envolve (a) segmentação da cobertura vegetal ou reconhecimento de entidades, (b) seleção de unidades amostrais nos segmentos reconhecidos, (c) decisão sobre tamanho e forma da unidade, e (d) decisão sobre o que registrar após as unidades serem estabelecidas. A segmentação sempre é subjetiva. Quando os segmentos em (a) são homogêneos, como em Braun-Blanquet (1964), apenas uma unidade amostral de tamanho adequado (área mínima fitossociológica) pode ser suficiente para descrevê-los. A descrição e análise pode ou não confirmar os padrões (segmentos) definidos a priori, configurando um processo de sucessivas aproximações (Poore 1956, 1962, Orlóci 1991). Quando os segmentos reconhecidos são maiores, e por consequência mais heterogêneos, não é possível descrevê-los com apenas uma unidade amostral, sendo necessária uma seleção objetiva (aleatória ou sistemática) das unidades amostrais. Assim, a segmentação pode ser interpretada como o processo mesmo de delimitar comunidades vegetais.

Portanto, a descrição de comunidades vegetais pode envolver ao mesmo tempo uma amostragem objetivando estudar padrões e outra objetivando a estimativa de quantidades. A seleção e delimitação de um conjunto de comunidades vegetais em uma área de vegetação é uma amostragem com vistas a estudar padrões. Essas comunidades vegetais serão descritas, e para tanto pode ser também necessária uma amostragem em cada comunidade para estimar quantidades das espécies componentes, especialmente quando o tamanho da comunidade vegetal é tal que um censo da sua totalidade não é possível. É importante não confundir esses dois níveis da amostragem, os quais terão objetivos e procedimentos distintos.

O número de unidades amostrais

Um problema relevante na amostragem, e sempre presente, é decidir o número de unidades amostrais a usar (o *tamanho da amostra*), pois dele depende em grande parte a quantidade de trabalho envolvido na descrição do conjunto de unidades amostrais. A interpretação da amostragem como um processo de sucessivas aproximações tem precedentes em Greig-Smith (1957) para a estimativa de atributos simples e em Orlóci & Pillar (1989) para o estudo de padrões. Nessa abordagem o estado de um dado atributo obtido a partir da amostra evolui e atinge estabilidade na medida em que se aumenta o número de unidades amostrais na amostra. O tamanho ótimo de amostra é aquele no qual o atributo simples ou complexo de interesse começa a ter estabilidade, ou seja, quando o fato de agregar-se novas unidades amostrais à amostra resulta em alterações relativamente menores no valor do atributo considerado. Assim, se por um lado o objetivo é estimar a média de alguma variável, o tamanho ótimo da amostra será aquele em que a média da amostra atinge estabilidade. Um exemplo simples é o caso em que a média e o desvio padrão na amostra é monitorado para tamanhos sucessivos de amostra (Figuras 1 e 2).

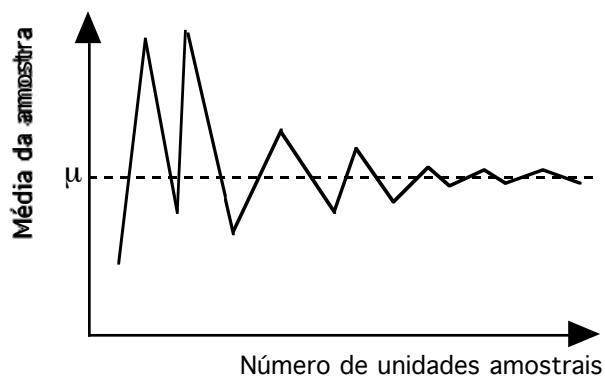


Figura 1. A estimativa do atributo será mais acurada quanto mais próximo do estado verdadeiro do universo amostral for o estado inferido via amostragem.

Um exemplo dessa abordagem para determinação de tamanho de amostra é a curva "número de espécies x número de unidades amostrais", muito usada em ecologia de comunidades para, entre outros objetivos, indicar suficiência de amostragem; o atributo considerado é o número de espécies. A curva "número de espécies x tamanho da unidade amostral", usada para determinar a área mínima fitossociológica, é um caso análogo; o processo nesse caso pode ser entendido como uma agregação de novas unidades amostrais sistemática e contiguamente às que já estão na amostra. Quaisquer outros atributos, simples ou complexos (e.g., medidas de diversidade), poderiam também ser considerados nessas curvas.

Por outro lado, se o objetivo é estudar padrões e suas relações, dependendo do tipo de dado, haverá inúmeras maneiras de representar numericamente a estrutura de dados na amostra e suas relações. O caso mais simples é se apenas duas variáveis descrevem as unidades amostrais. O exemplo na Figura 3 foi obtido de um estudo que objetivou investigar em uma área plantada com eucalipto o comportamento da vegetação campestre natural remanescente frente ao incremento do sombreamento. Observa-se uma estabilização da correlação entre a cobertura com *Paspalum notatum* e o grau de sombreamento a partir de 7 unidades amostrais, que seria então o tamanho ótimo de amostra para essas variáveis.

O processo é o mesmo quando mais de duas variáveis estão envolvidas. Por exemplo, quando a descrição da vegetação é através da avaliação das quantidades das várias espécies e de variáveis de ambiente nas mesmas unidades amostrais, e o objetivo é revelar a relação entre a vegetação e as variáveis de ambiente. O exemplo na Figura 4, que se refere à área com eucalipto já mencionada, mostra uma estabilização da correlação entre vegetação e sombreamento ao redor de 11 unidades amostrais. Do mesmo modo, para mapear ou classificar a vegetação, o tamanho ótimo será aquele em que se estabiliza a função matemática definida para avaliar a estrutura de grupos nos dados.

Tabela 2. Dados de pH do solo obtidos em 15 unidades amostrais tomadas ao acaso, no Parque Estadual de Vassununga, Santa Rita do Passa Quatro, São Paulo (Martins 1991:105). A terceira e quarta linhas indicam a média e o desvio padrão calculados neste trabalho para cada um dos tamanhos crescentes da amostra, de 1 a 15 unidades amostrais, mostrados na Figura 2.

n°	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
pH	5.3	5.6	6.4	6.6	5.7	7.3	7.1	7	5.6	7.2	6.6	5.9	7.3	6.7	6.3
\bar{x}	5.3	5.45	5.77	5.97	5.92	6.15	6.29	6.37	6.29	6.38	6.4	6.36	6.43	6.45	6.44
s	0	0.21	0.57	0.62	0.55	0.75	0.77	0.76	0.76	0.77	0.73	0.71	0.73	0.71	0.68

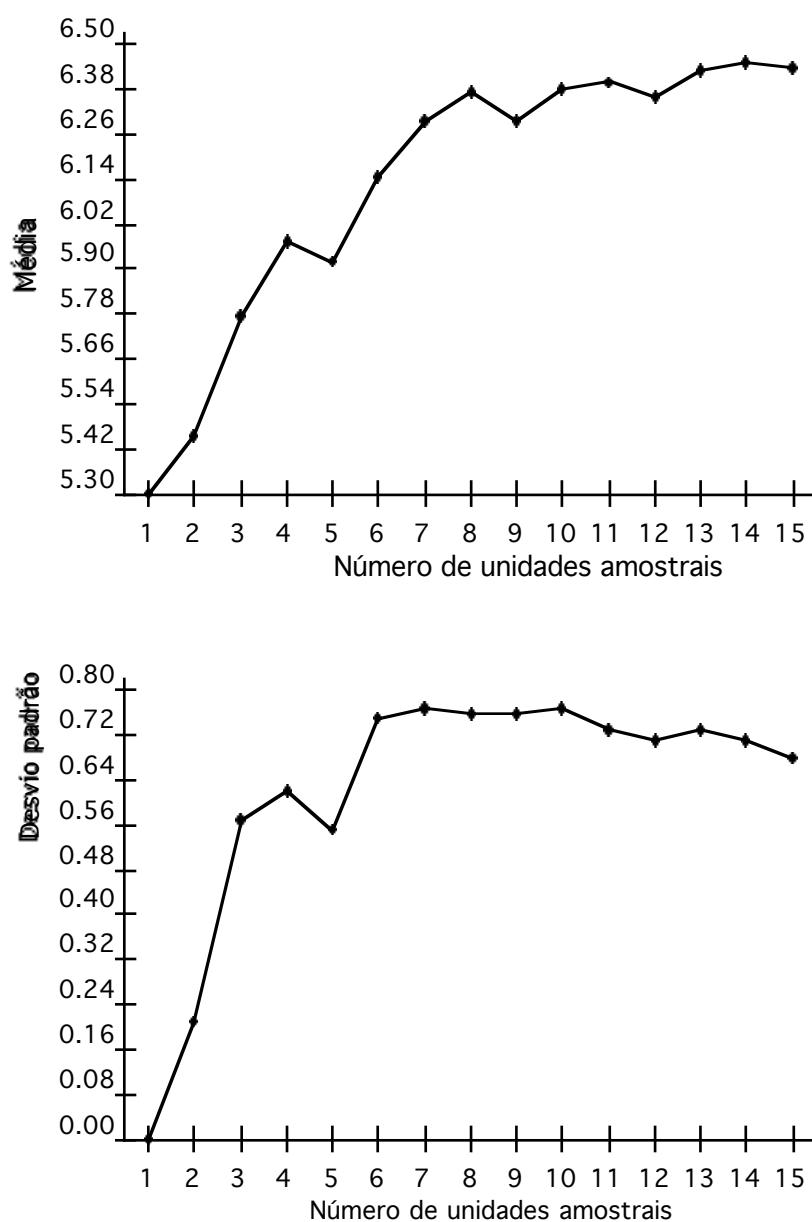


Figura 2. Média e desvio padrão para um tamanho crescente da amostra até 15 unidades amostrais, tomadas ao acaso, no Parque Estadual de Vassununga, Santa Rita do Passa Quatro, São Paulo (dados de pH de Martins 1991:105). Note a estabilização relativa da média e desvio padrão a partir de 8 e 6 unidades amostrais respectivamente.

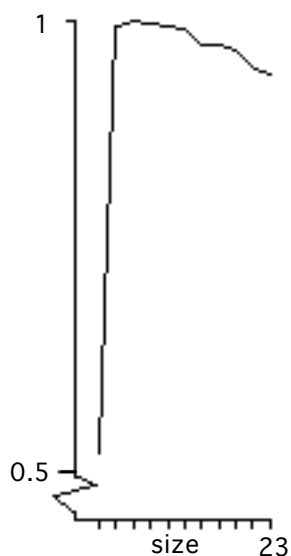


Figura 3. Evolução da correlação ($1-r^2$ no eixo vertical) matricial entre as distâncias baseadas na abundância-cobertura de *Paspalum notatum* e as distâncias baseadas na interceptação da radiação solar para uma amostra de tamanho crescente. Dados obtidos em quadros de 0,5 m de lado, sobre a vegetação campestre remanescente nas entrelinhas de um plantio de eucalipto, Horto Tipuana, Arroio dos Ratos, RS (Pillar, Boldrini & Lange 1992). A ordem de inclusão dos quadros (unidades amostrais) na amostra é 1 8 15 16 20 24 4 11 17 22 2 6 9 13 3 5 7 10 12 14 18 21 23. Note a estabilização relativa da correlação a partir de 7 unidades amostrais, que seria o tamanho ótimo de amostra para essas variáveis. A área está cultivada com eucalipto desde há 3 anos. O objetivo do estudo é investigar o comportamento da vegetação campestre natural remanescente nas entre-linhas frente ao incremento do sombreamento pelo eucalipto a partir do aceiro até alguns metros mata adentro. As unidades amostrais foram locados ao longo de uma transecção a partir do aceiro e que penetra na mata em direção S (15 quadros) e N (8 quadros), a cada 1 m metro nos quadros mais próximos do aceiro, e a cada 1,5 m nos quadros mais sombreados. A radiação total é uma média da radiação medida na altura de 20 cm do solo durante 3 dias no mês de agosto/92 a cada 1/2 hora no primeiro dia e a cada 1 hora nos outros dias. O grau de interceptação foi calculado em relação à radiação medida simultaneamente na ausência de sombreamento. Os dados de radiação foram obtidos por Homero Bergamaschi e colaboradores.

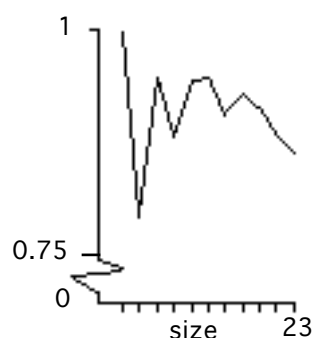


Figura 4. Evolução da correlação ($1-r^2$ no eixo vertical) matricial entre as distâncias baseadas na abundância-cobertura de 63 espécies e as distâncias baseadas na interceptação da radiação solar para uma amostra de tamanho crescente. Dados obtidos no Horto Tipuana de acordo com a metodologia descrita na legenda da Figura 3. A ordem de inclusão dos quadros (unidades amostrais) na amostra é 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 20 21 22 23 24. Note a estabilização relativa da correlação a partir de 11 unidades amostrais, que seria o tamanho ótimo de amostra nesse caso.

Bibliografia

- Braun-Blanquet, J. 1928-1964. *Fitosociologia; bases para el estudio de las comunidades vegetales*. Trad. da 3.ed.rev.aum. Blume, Madrid, 1979. 820 p.
- Greig-Smith, P. 1957. *Quantitative Plant Ecology*. Butterworths, London. 198pp.
- Kenkel, N. C., P. Juhász-Nagy & J. Podani. 1989. On sampling procedures in population and community ecology. *Vegetatio* 83: 195-207.
- Martins, F. R. 1991. *Estrutura de uma Floresta Mesófila*. Editora da UNICAMP, Campinas. 246 pp.
- Mueller-Dombois, D. & H. Ellenberg. 1974. *Aims and Methods of Vegetation Ecology*. Wiley, New York. 547 p.
- Orlóci, L. 1991. Poorean approximation and fisherian inference in bioenvironmental analysis. *Advances in Ecology 1*: 65-71.
- Orlóci, L. & Pillar, V. D. P. 1989. On sample size optimality in ecosystem survey. *Biometrie-Praximetrie* 29:173-184.
- Pillar, V. D. P., Boldrini, I. O. & Lange, O. 1992. Padrões de variação da vegetação campestre sob eucalipto. Departamento de Plantas Forrageiras e Agrometeorologia, UFRGS. (manuscrito)
- Podani, J., T. Czárán & S. Bartha. 1993. Pattern, area and diversity: the importance of spatial scale in species assemblages. *Abstracta Botanica* 17: 37-51.
- Poore, M. E. D. 1956. The use of phytosociological methods in ecological investigations. IV. General discussion of phytosociological problems. *Journal of Ecology* 44: 28-50.
- Poore, M. E. D. 1962. The method of successive approximation in descriptive ecology. *Advances in Ecological Research 1*: 35-69, Academic Press, New York.