Crescimento denso-independente com estocasticidade ambiental - Roteiro no EcoVirtual

No modelo de crescimento discreto independente de densidade projetamos o tamanho população multiplicando o seu tamanho inicial N_0 pela mesma taxa de crescimento λ em todas as t gerações:

$$N_t=N_0 \lambda^t \tag{eq1}$$

Com isso, estamos assumindo que a taxa de crescimento é a mesma a cada geração, o que é pouco realista. Condições e recursos variam ao longo do tempo, o que deve fazer λ flutuar. Se aceitamos isso, e chamando de λ_i os valores da taxa de crescimento a cada geração[1)], nosso modelo passa a ser:

$$N_t=N_0 \lambda_1 \lambda_2 \lambda_3 \ldots \lambda_t \ = \ N_0 \prod_{i=1}^t \lambda_i \tag{eq2}$$

Ainda assim, teremos uma taxa média de crescimento, que poderíamos estimar com a média das taxas observadas ao longo das gerações. Se o ambiente é pouco variável, as taxas observadas deveriam estar próximas dessa média a maior parte do tempo, ou seja, os valores λ_i teriam pouca variância[2)]. Se há muita variação ambiental, isso deve se refletir em taxas de crescimento mais variáveis. Chamamos de *estocasticidade ambiental* a incerteza no valor de taxas demográficas, devido às flutuações nas condições e recursos experienciadas por todos os indivíduos da população.

Simulação de estocasticidade ambiental

Para prosseguir você deve ter o ambiente **R** com os pacotes **Rcmdr** e **Ecovirtual** instalados e carregados. Se você não tem e não sabe como ter, consulte a página de Instalação.

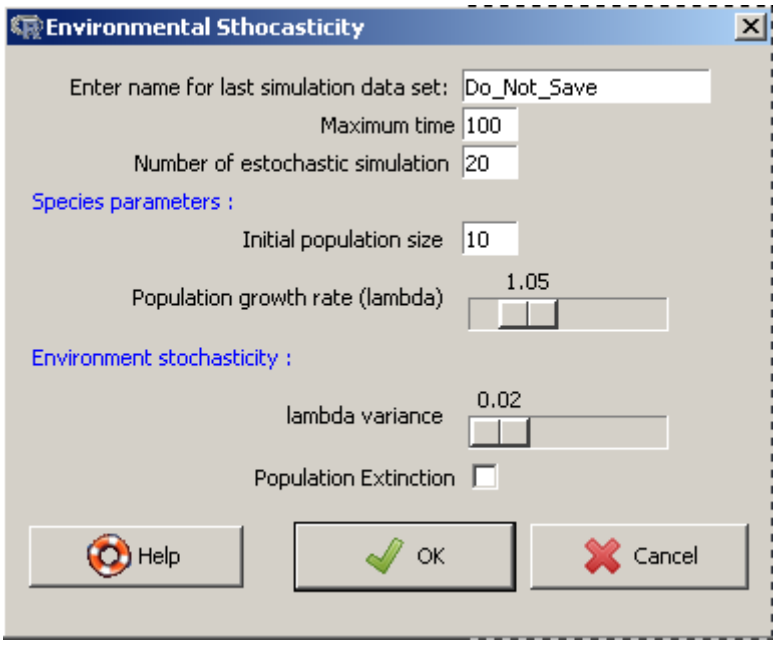
Caso já tenha o R e pacotes instalados

Carregue o pacote principal ***RcmdrPlugin.EcoVirtual*** pelo menu do R **Pacotes** > **Carregar Pacotes**, ou pela linha de comando com o código:

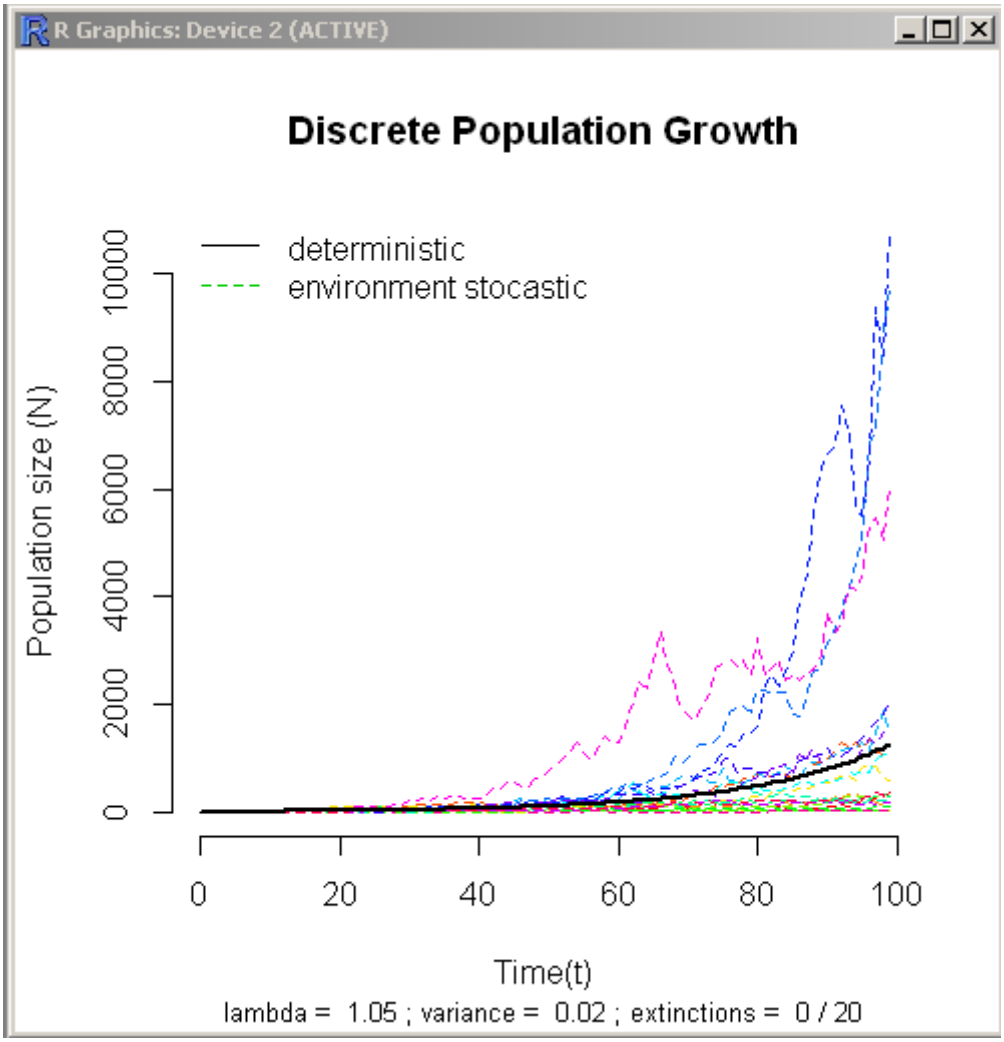
```
library("RcmdrPlugin.EcoVirtual")
```

De que maneira a estocasticidade ambiental afeta as projeções de tamanho populacional? Vamos responder isso simulando o modelo de crescimento discreto (equação $\ref{eq2}$), com valores de λ_i que mudam a cada geração. Escolhemos este modelo por razões didáticas e de implementação computacional, mas as conclusões são as mesmas para modelos em tempo contínuo.

Na janela do Rcmdr clique na opção de menu *EcoVirtual*, e depois em *One population* e em seguida em *Environmental Stochasticity*. Duas janelas irão se abrir, uma de opções como esta:



e outra com um gráfico como este:



Arranje as janelas de modo que o gráfico esteja sempre visível, mesmo quando você alterar valores na janela de opções.

Parâmetros da função

Abaixo os parâmetros do nosso modelo:

Opção	parâmetro	definição
Enter name for last simulation data set	objeto no R	nome para salvar os resultados da simulação em um objeto no R
Maximum time	txmax	número de intervalos de tempo
Number of stochastic simulations	npop	número de populações a simular
Initial population size	N0	tamanho inicial das populações
Population growth rate	lambda	taxa de crescimento média
lambda variance	varr	variância da taxa de crescimento
Population extinction	ext	populações extinguem-se caso cheguem a tamanhos menores que um? TRUE para sim; FALSE para não

O que é o gráfico?

Cada linha colorida é a projeção no tempo do tamanho de uma população de acordo com o modelo ($\ref{eq2}$). Todas as populações têm o mesmo valor médio de taxa de crescimento e mesmo tamanho inicial. As projeções diferem porque a cada intervalo de tempo é sorteado um valor da taxa de crescimento independente para cada população.

Os sorteios são feitos de uma distribuição de probabilidades chamada lognormal, que garante que as taxas serão sempre positivas, como exige o modelo. Você pode mudar a média e variância da lognormal usada no sorteio (veja a seguir). Quanto maior a estocasticidade ambiental, maior a variância da taxas de crescimento.

A linha preta indica a projeção de uma população sob uma taxa constante igual à da média da distribuição de taxas. Apesar do modelo ser discreto, representamos as projeções com linhas para facilitar seguir a trajetória das populações no tempo.

Familiarizando-se com as simulações

Nosso objetivo é entender os efeitos da estocasticidade demográfica sobre o crescimento discreto independente da densidade. Vamos começar avaliando o comportamento de uma população por vez:

1. Altere o valor de `Maximum time (tmax)` para 10;
2. Altere o valor da opção `Number of stochastic simulations (npop)` para 1;
3. Altere o valor de `Population growth rate (lambda)` para 2;

1. Repita algumas vezes a operação e observe o gráfico.
2. Repita com valores mais altos de estocasticidade ambiental: para isso, aumente o valor em `lambda variance(varr)` e observe o gráfico.

A cada vez que você roda a função, o **EcoVirtual** sorteia uma valor de λ para cada intervalo de tempo e os utiliza para projetar o tamanho da população ao longo do tempo, conforme a equação $\ref{eq2}$.

Os sorteios são tomados da mesma distribuição de probabilidade, com a média e variância especificadas. Os sorteios são *independentes*, isto é, o valor em um momento não afeta o valor sorteado para os intervalos seguintes. Por isso dizemos que as taxas sorteadas são variáveis aleatórias independentes mas de igual distribuição de probabilidades.

Perguntas

1. O que é preciso para simular ausência de estocasticidade ambiental?
2. Qual o comportamento da projeção na ausência de estocasticidade ambiental? Verifique sua previsão com o **EcoVirtual**

Média de muitas projeções

Uma maneira de lidar com processos variáveis é descrever seu comportamento médio. Por exemplo,

aqui está um gráfico em tempo real do congestionamento na cidade de São Paulo ao longo das horas do dia de hoje, em comparação com as médias históricas dos menores e maiores valores de cada horário. O percentual de vias congestionadas varia a cada dia, mas isso não o torna completamente imprevisível. Se acumulamos dados de muitos dias podemos ter uma ideia do que esperar, e mesmo avaliar se estamos em um dia atípico. Este valor esperado, que designamos $E[X]$, é a média teórica de uma variável aleatória X [3)], no caso o percentual da extensão de vias que estão congestionadas numa dada hora.

Em nossas simulações, os tamanhos populacionais são resultado da multiplicação de taxas de crescimento sorteadas sempre de uma mesma distribuição de probabilidades. Se esses produtos tiverem um valor esperado definido poderíamos prever o comportamento das populações, mesmo com a incerteza da estocasticidade ambiental. Será possível?

Vamos avaliar isso projetando várias populações independentes, para ter muitos valores de tamanhos populacionais a cada tempo. Para isso entre os seguintes valores na função:

```
# grave o resultado da função com o nome sim1

tmax = 51
npop = 1000
N0 = 10
lambda = 1.1
varr = 0.03
ext = FALSE
```

Rode a função e verá um gráfico com a evolução independente de 1000 populações ao longo do tempo, segundo o modelo ($\ref{eq2}$), para um λ médio de 1,1 com variância de 0,03.

Ao colocar um nome para o resultado da simulação, você gravou as projeções das populações em cada tempo na memória do ambiente R. É uma tabela de dados em que as colunas são os tamanhos populacionais e as linhas os intervalos de tempo. É o que precisamos para calcular o tamanho médio das populações a cada tempo!

Usando o Rcmdr para rodar códigos

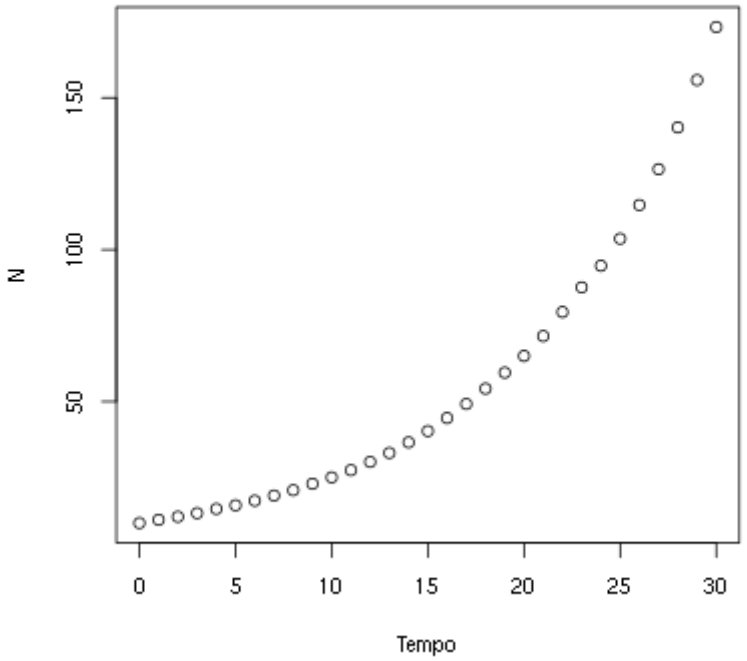
O Rcmdr tem 3 janelas sobrepostas. A superior recebe o script ou código que se quer rodar, no formato de texto. Para submeter linhas ou parte do código para o R é preciso selecionar o texto ou colocar o cursor na linha e clicar em submeter. Há três tipos de retorno a uma submissão de código do R: resultado alfanumérico de operações, gráfico ou mensagem de erro. Qualquer resultado de operações que possa ser representado em caracteres será lançado na janela intermediária chamada de `Output`. Gráficos serão mostrados em um janela própria. A janela inferior é reservada para mensagens de erro. Fique atento, quando há mensagem de erro significa que seu código não foi avaliado, chame um monitor.

Vamos fazer os cálculos e um gráfico dos tamanhos médios em função do tempo no próprio R. Copie

os seguintes comandos do R na janela `script` do Rcmdr [4]:

```
## Elimina as duas primeiras colunas da tabela (tempo e valor
deterministico)
sim1b <- sim1[,-(1:2)]

## Calcula a media para cada tempo
medias=apply(sim1b,1,mean)

## Vetor de tempo de zero ao maximo
tempo <- length(medias)-1

## Grafico com medias
plot(0:tempo,medias, xlab="Tempo", ylab="N")
```

e agora rode estes comandos no R. Se tudo deu certo a janela gráfica vai mostrar uma figura como esta:



As médias parecem ter um crescimento geométrico, similar ao modelo de crescimento discreto sem estocasticidade. Vamos verificar isso para os cinco primeiros valores da média das projeções. Para vê-los, copie o comando abaixo na janela `Script` e rode-o:

```
# Visualizando as médias das projecoes de t=0 a t=5.

medias[1:6]
```

Agora calcule as projeções para os mesmo tempos usando o λ médio. O comando abaixo

executa essa conta em único comando. O codigo `0:5` cria o vetor de valores `0, 1, 2, 3, 4, 5` ...

```
## lembre-se que a populacional que simulou tinha 10 indivíduos inicialmente e o lambda de ''1.1''

10*1.1^(0:5)
```

Para ajudar na comparação entre estas duas projeções, podemos adicionar a que usa o λ médio ao gráfico com os comandos

```
points(0:tempo, 10*1.1^(0:tempo), pch=19, col="blue", cex=0.5)

legend("topleft", c("medias observadas", "projecao com lambda medio"),
        pch=c(1,19), col=c("black", "blue"))
```

Pergunta

Dados os resultados que você obteve e o modelo de crescimento discreto com estocasticidade ambiental da equação $\ref{eq2}$:

$$N_t \ = \ N_0 \prod_{i=1}^t \lambda_i$$

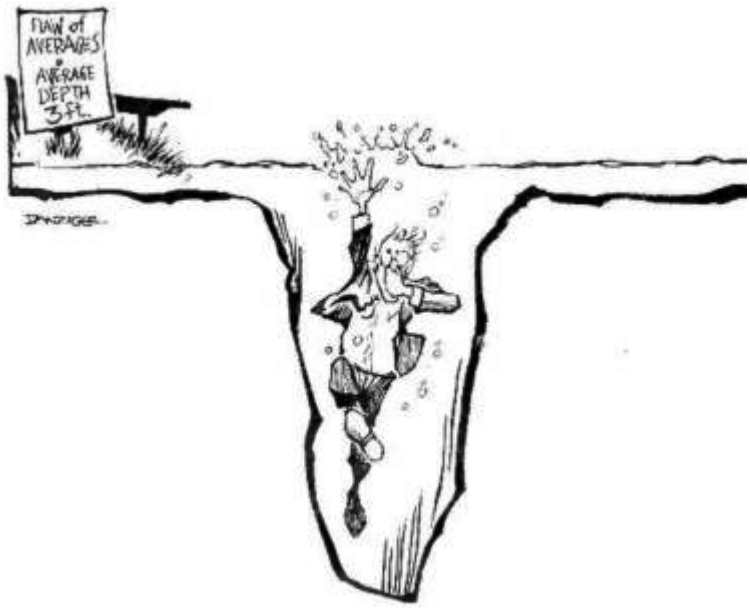
e se definimos a média dos valores de λ_i como

$$\bar \lambda \ = \ E[\lambda_i]$$

proponha a equação para o valor esperado do tamanho populacional $E[N_t]$ em função de $\bar \lambda$.

Variância de muitas projeções

A seção anterior mostra que mesmo com estocasticidade ambiental as projeções em média são as mesmas do modelo sem estocasticidade.



É um resultado reconfortante mas, como toda síntese, a média pode ser traiçoeira, como diria o estatístico que se afogou no lago com profundidade média de 90 cm. A figura desse pobre homem da ciência, ao lado, é um *link* para um ótimo *post* com mais exemplos de médias enganadoras.

A informação que falta é quanta variação há nos dados. Uma das maneiras de medir isso é a *variância*, que expressa o desvio esperado dos dados em relação à média. Para uma variável aleatória X a variância teórica é:

$$VAR[X] \ = \ E[(X - E[X])^2]$$

Note que a variância é a média do desvio em relação à média. Este desvio está elevado ao quadrado para ter algumas propriedades convenientes[5)]. Por isso a variância também é chamada *desvio quadrático médio* ou *desvio quadrático esperado*.

Mas há uma inconveniência de expressar a variação com desvios quadráticos: as unidades da variância são o quadrado das unidades da medida original. Se a média estiver em centímetros, a variância desta média estará em centímetros quadrados 🤔. Isso é resolvido tomando-se a raiz quadrada da variância, que chamamos *desvio-padrão*.

Bonito, mas como calculo isso na prática?

A expressão da variância acima é uma definição teórica. Na prática não sabemos valores esperados teóricos, mas podemos estimá-los. Fazemos isso de um conjunto de medidas, supondo que estas medidas são uma amostra representativa do processo cujo valor esperado queremos estimar. Assim, em uma amostra de n medidas estimamos a variância com a expressão:

$$s^2 \ = \ \sum_{i=1}^n \frac{(x_i - \bar x)^2}{n-1}$$

onde $\bar x$ é a estimativa da média

$$\bar x \ = \ \frac{1}{n}\sum_{i=1}^n x_i $$

e n o tamanho da amostra (seu número de elementos medidos). A estimativa do desvio-padrão é obtida com a raiz quadrada de s^2:

$$s \ = \ \sqrt{\sum_{i=1}^n\frac{(x_i - \bar x)^2}{n-1}} $$

Desvio-padrão das projeções com estocasticidade ambiental

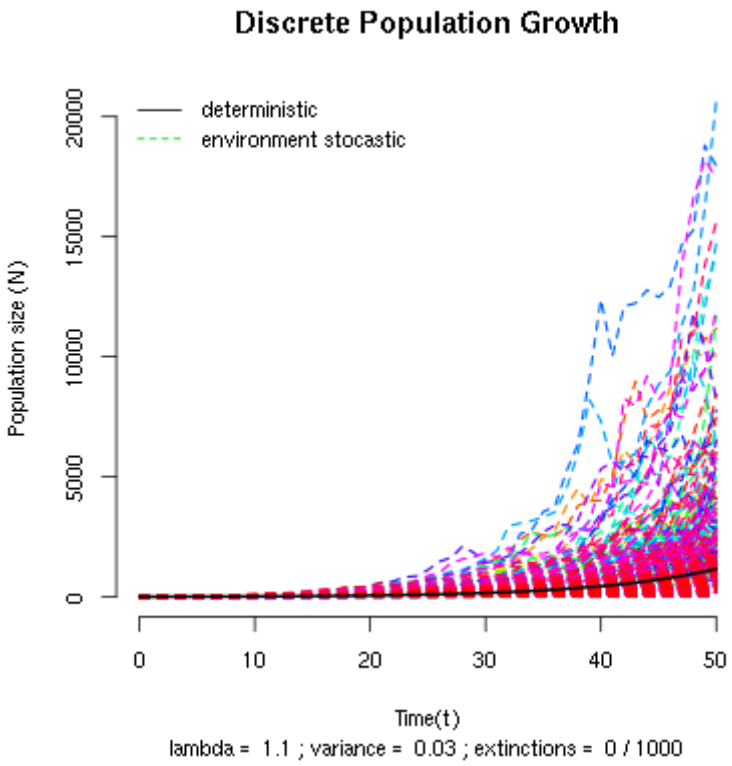
Com isso podemos avaliar como as projeções de nosso modelo com estocasticidade variaram ao longo do tempo. Vamos usar as projeções da última simulação, que já guardados em uma tabela na memória do ambiente R. Para criar um gráfico da média e desvio-padrão em função do tempo copie e rode os comandos abaixo:

```
## calcula os desvios-padrão para cada tempo
desvios <- apply(sim1b,1,sd)

## Medias em funcao do tempo
plot(0:tempo,medias, xlab="Tempo", ylab="N")
```

```
#3 Acrescenta desvios-padrao ao grafico
points(0:tempo,desvios, xlab="Tempo", ylab="N", col="blue")

## Adiciona legenda ao grafico
legend("topleft", c("medias", "desvios-padrão"),
        pch=c(1,1), col=c("black", "blue"))
```

O desvio-padrão também cresce exponecialmente com o tempo! Este crescimento pode ser até mais rápido que o crescimento do tamanho populacional esperado, como para a simulação que usamos aqui.

Tomando o valor do desvio-padrão como uma medida de incerteza de nossa expectativa, podemos dizer que projeções a longo prazo são extremamente imprecisas, apesar de terem um valor esperado conhecido. Agora que você já sabe isso, rode de novo a simulação e veja o gráfico, que deve ser parecido com este



Note como a amplitude das projeções se abre como um funil à medida que o tempo passa. Isso acontece porque as trajetórias divergem com a multiplicação de taxas variáveis, mesmo que as populações partam do mesmo valor inicial.

Risco de Extinção

A variância e o desvio-padrão são médias (dos desvios) e, portanto, também podem não caracterizar completamente a estrutura de variação dos dados. Considere este dois conjuntos de valores[6]:

Conjunto A										
4	5	6	7	8	9	10	11	12	13	14	15

Conjunto B											
8	8	8	8	8	8	8	8	8	8	8	19

para os dois conjuntos a média é exatamente 9 e a variância exatamente 11. O que escapa aos dois índices é a distribuição dos valores em relação à média, que é bem mais *assimétrica* no conjunto B, em que apenas um valor é maior que a média.

Agora olhe de novo o gráfico das projeções da seção anterior e avalie a simetria da distribuição dos valores projetados em relação à média, indicada pela linha preta. Para ajudar, execute o código do R abaixo para fazer histogramas das projeções nos tempos 5, 10, 20 e 30:

```
## valor maximo para definir a escala do histograma
sim1b.m <- (max(sim1b[c(6,11,21,31),]))

##define uma janela para 4 gráficos
par(mfrow=c(2,2))

##Cria os 4 histogramas
hist(sim1b[6,], xlab="N", ylab="Frequencia", main="T=5",
     breaks=seq(0,max(sim1b[6,]+50),by=50), xlim=c(0,sim1b.m))
hist(sim1b[11,], xlab="", ylab="", main="T=10",
     breaks=seq(0,max(sim1b[11,]+50),by=50), xlim=c(0,sim1b.m))
hist(sim1b[21,], xlab="", ylab="", main="T=20",
     breaks=seq(0,max(sim1b[21,]+50),by=50), xlim=c(0,sim1b.m))
hist(sim1b[31,], xlab="", ylab="", main="T=30",
     breaks=seq(0,max(sim1b[31,]+50),by=50), xlim=c(0,sim1b.m))
## Volta a janela para 1 gráfico
par(mfrow=c(1,1))
```

As distribuições das projeções tornam-se cada vez mais assimétricas com o passar do tempo. Isso acontece porque poucas populações crescem muito, enquanto a maioria permanece em tamanhos pequenos. Dependendo da média e variância das taxas de crescimento, muitas populações podem até diminuir de tamanho.

Assim, mesmo com taxa média de crescimento maior que um ($\bar \lambda>1$) a estocasticidade ambiental pode fazer com que a maioria das populações não cresça, ou mesmo diminua! Vamos verificar isso contando quantas populações passam do tamanho inicial em função do tempo em uma simulação em que λ tem uma variância alta.

Execute as simulações abaixo na função de estocasticidade ambiental:

```
# salve os resultados da simulação no objeto 'sim2'
tmax = 51
npop = 1000
N0 = 10
lambda = 1.05
varr = 0.2
ext = FALSE
```

Os resultados estão guardados em uma nova tabela na memória do R, chamada sim2. Com ela podemos calcular a proporção das 1000 populações simuladas que tiveram tamanhos maiores do que

$N_0=10$ a cada tempo. Para isso, copie e execute os comandos abaixo:

```
maiorN0 <- apply(sim2[,-(1:2)],1,
                 function(x)sum(x>10)/length(x))
plot(sim2[-1,1],maiorN0[-1], xlab="Tempo",
     ylab="Proporcao projecoes > N0")
```

A proporção de projeções maiores do que N_0 cai com o tempo. Isso indica que a probabilidade de uma população efetivamente crescer é cada vez menor, mesmo que em média haja crescimento .

Este aparente paradoxo se explica pela forte assimetria das distribuições das projeções populacionais: a maioria das populações está abaixo da média, o que se acentua com o tempo [7]. E mais: se a variância ambiental [8] é grande, cada vez menos populações estarão acima de um valor abritrário qualquer, como o tamanho inicial da população.

Pergunta

Na função de estocasticidade ambiental, o parâmetro ext se marcado (ext=TRUE) executas as simulações com $N_{min}=1$, ou seja, faz com que toda população que caia abaixo de um extinga-se[9]. O número de populações que se extinguiram ao final da simulação é indicado abaixo do eixo X. Use esta opção para descobrir:

1. Utilize o tamanho inicial da população como $N_0 = 100$ e $N_0 = 1000$ para invetigar se o tamanho inicial das populações modificam o resultado das questões abaixo.
2. Qual o efeito do aumento de estocasticidade ambiental sobre a probabilidade de extinção?
3. Dado um certo nível de estocasticidade, qual o efeito de mais tempo de simulação sobre a probabilidade de extinção?
4. Deduza uma consequência desses resultados para a conservação de populações ameaçadas.

Uma boa estimativa da probabilidade de um evento é a proporção com que ele acontece em um grande número de tentativas. Assim, use um grande número de populações nas suas simulações. Em alguns casos o gráfico vai ficar muito cheio de linhas, mas o número que interessa está abaixo do eixo X, e continuará legível.

Para saber mais

- On population growth in a randomly varying environment: neste artigo clássico, Richard Lewontin e Dan Cohen chamaram a atenção dos biólogos para as propriedades surpreendentes dos modelos que exploramos aqui.
- tutorial em R de multiplicação de variáveis aleatórias: veja como o modelo que usamos demonstra que o resultado desta multiplicação tende a uma distribuição lognormal. Da disciplina de de modelagem probabilística da pós-graduação em ecologia da USP.
- Life is log-normal : argumentação provocadora a favor do uso da log-normal ao invés da normal,

com link para o artigo e uma demonstração animada.

Rcmdr, uma população, crescimento exponencial, tempo discreto, tempo contínuo, estocasticidade ambiental

1)

$i=1 , 2, 3, \ldots t$

2)

mais adiante neste roteiro definimos precisamente variância. Até lá basta a noção de que ela é uma medida da variação dos dados

3)

ou seja, uma variável que pode assumir diferentes valores

4)

as linhas precedidas por um ou mais “#” não são executadas pelo R. Elas são comentários sobre o que o comando na linha seguinte faz

5)

para saber sobre estas propriedades e como estimar a variância de uma amostra de dados veja aqui

6)

estes dados fazem parte de um conjunto criado pelo estatístico Francis Anscombe para ilustrar algumas armadilhas dos índices estatísticos. Veja mais aqui

7)

é possível provar mais precisamente este resultado. Com o passar do tempo a distribuição de projeções deste modelo tende a uma distribuição lognormal com variância crescente, o que a torna cada vez mais assimétrica

8)

neste modelo representada pela variância de λ

9)

$N_{min}=1$ parece uma escolha natural, mas se em nosso modelo N_t é um valor de densidade e não de número absoluto de indivíduos outro valores de N_{min} são válidos.

From:
http://ecovirtual.ib.usp.br/ -

Permanent link:
http://ecovirtual.ib.usp.br/doku.php?id=ecovirt:roteiro:den_ind:di_earcmdr

Last update: **2016/05/10 07:19**