2022

Informe de
ASESORAMIENTO
y TRANSFERENCIA

046-22

Generación de valores de reclutamiento con media una función de reclutamiento y distribución log-normal

Aubone Aníbal

Resumen

Predecir el reclutamiento es una aspiración de los investigadores pesqueros de hace mucho tiempo. Sin embargo, es difícil por las distintas variables biológicas y factores abióticos que lo afectan. Para algunos modelos de dinámica poblacional el uso de funciones de reclutamiento es requerido como alternativa para disminuir el número de parámetros a estimar. Pero esto no deja de tener implicaciones importantes. En este trabajo se analiza el uso de funciones de reclutamiento (que relacionan los reclutas con la biomasa de hembras desovante que les dio origen), los supuestos subyacentes y la introducción de aleatoriedad en los desvíos del reclutamiento de los valores medios, en los modelos de dinámica de poblaciones estructuradas por edades. El supuesto de distribución log-normal de errores se plantea formalmente y se discute sobre la manera de considerarlo en el modelo de dinámica poblacional. Se discute sobre los efectos de los supuestos no validables, que conlleva tanto el uso de funciones de reclutamiento, como de introducir la aleatoriedad en los desvíos de los reclutamientos de la misma y el impacto sobre las inferencias a largo plazo.

Ministerio de Agricultura, Ganadería y Pesca
Argentina

INIDEP
INSTITUTO NACIONAL DE INVESTIGACIÓN Y DESARROLLO PESQUERO

Generación de valores de reclutamiento con media una función de reclutamiento y distribución log-normal

Aubone, Aníbal

Instituto Nacional de Investigación y Desarrollo Pesquero
Universidad Nacional de Mar del Plata

Resumen

Predecir el reclutamiento es una aspiración de los investigadores pesqueros de hace mucho tiempo. Sin embargo, es difícil por las distintas variables biológicas y factores abióticos que lo afectan. Para algunos modelos de dinámica poblacional el uso de funciones de reclutamiento es requerido como alternativa para disminuir el número de parámetros a estimar. Pero esto no deja de tener implicaciones importantes. Se analiza el uso de funciones de reclutamiento (que relacionan los reclutas con la biomasa de hembras desovante que les dio origen), los supuestos subyacentes y la introducción de aleatoriedad en los desvíos del reclutamiento de los valores medios, en los modelos de dinámica de poblaciones estructuradas por edades. El supuesto de distribución log-normal de errores se plantea formalmente y se discute sobre la manera de considerarlo en el modelo de dinámica poblacional. Se discute sobre los efectos de los supuestos no validables, que conlleva tanto el uso de funciones de reclutamiento, como de introducir la aleatoriedad en los desvíos de los reclutamientos de la misma y el impacto sobre las inferencias a largo plazo.

Palabras Clave
Generación de reclutamientos, distribución log-normal, funciones de reclutamiento

Introducción

Predecir el reclutamiento ha sido un objetivo de la investigación pesquera por más de cien años. Desafortunadamente, no se ha logrado el éxito en términos generales, debido a múltiples mecanismos que afectan al reclutamiento en las etapas primeras de vida de los individuos, y el pobre conocimiento de los mismos. Lo que se plantea es la relación entre el reclutamiento y los padres que les dieron origen; pero el reclutamiento no sólo depende de la abundancia de padres, sino que también de otros factores como el crecimiento, mortalidad, estructura poblacional, fertilidades, alimento disponible, medio ambiente, depredación, etc.; procesos complejos que requieren de investigaciones detalladas, que muchas veces no están disponibles. Entonces, lo que suele utilizarse es una relación simplificada entre el reclutamiento (tal vez de individuos de edad 1) y la biomasa de hembras desovantes o a veces la biomasa reproductiva que les dio origen.

En muchas situaciones de evaluación de un recurso pesquero, es necesario establecer condiciones sobre el reclutamiento, para poder realizar las estimaciones de los parámetros de los modelos de dinámica poblacional, estructurados por edades. Es importante que estas condiciones incluyan un índice de reclutamiento que permita ajustar la variabilidad del mismo, pero esto muchas veces no es posible por limitaciones de datos. Modelos secuenciales con el cálculo de abundancias para retrocediendo en el tiempo no suelen requerir supuestos o condicionamientos sobre el reclutamiento (por ejemplo, el Análisis de Poblaciones Virtuales). En modelos secuenciales de la familia ASPM (*Age Structured Production Model*), que utilizan patrones de selección para distribuir la captura total por edad, el cálculo de abundancias por edad se realiza con avance en el tiempo y en su versión básica, los reclutamientos son parámetros del modelo (Aubone, 2015 b). Esto genera una sobre-parametrización que se agrava con la incorporación de más tiempos, y complica las estimaciones de los parámetros. En estos casos, asumiendo una determinada función de reclutamiento (FR) puede disminuirse el número de parámetros involucrados. Pero la propia función de reclutamiento que se incorpore será un condicionamiento ya que se incorpora al modelo, en general, sin conocimiento biológico que la sustente. Existe mucha bibliografía donde se discute sobre la existencia de una FR, su validez en el tiempo, etc. (ver, por ejemplo, Aubone, 2015 a). Es muy importante tener en cuenta los efectos de largo plazo que se pueden producir, al utilizar una determinada FR como reclutamiento medio (Aubone, 2021). El problema surge si la elección de dicha función no es la representativa del reclutamiento medio. Existen procesos de reclutamiento que son compensatorios a pequeños valores de la biomasa de hembras desovantes y también procesos no compensatorios. Las conocidas FR de Beverton y Holt y de Ricker, indican procesos compensatorios fuertes del reclutamiento (la FR tiene derivadas primeras decrecientes, en un entorno del origen) y esto facilita la recuperación poblacional ante un aumento de la biomasa de hembras desovante. Pero podrían elegirse FR consistentes con procesos no compensatorios (por ejemplo, Beverton y Holt generalizada; con derivadas primeras crecientes en un entorno del origen). En estos procesos, la recuperación poblacional a pequeñas biomasas de hembras desovantes suele ser más costosa. Por lo tanto, no es inocuo el hecho de elegir una determinada FR (se asume que existe, pertenece a una familia de funciones determinada, es única y válida para todo tiempo).

Las FR más usuales son la función de Beverton y Holt, y la función de Ricker, pero como se dijo, no son las únicas utilizadas. Ambas FR determinan dinámicas particulares a largo plazo (Aubone, 2015 a, 2021), y es por ello que su uso e inferencia a largo plazo debe tomarse de manera precautoria. Además, ambas funciones tienen dos parámetros que requieren ser estimados. Sin embargo, existe un procedimiento que considera estados de equilibrio (y esto considera que la FR es la misma a largo plazo), por el cual se reduce el problema de estimar los parámetros a estimar un solo parámetro denominado el *steepness* h (Aubone, 2015 b, 2021). Sin embargo, será necesario estimar en el modelo de dinámica poblacional también el estado de equilibrio virgen (equilibrio supuesto, previamente al inicio de la explotación).

La posibilidad/capacidad de estimar los parámetros de una FR dependerá de la información subyacente en los datos y del modelo de dinámica poblacional planteado. Si no existen suficientes datos y datos consistentes con la FR considerada, no será posible lograr una estimación de los parámetros de la misma. Es necesario tener en cuenta las dificultades que esto puede significar para ciertos modelos de dinámica poblacional que dependen de la reducción del número de parámetros, y además de los supuestos asumidos como verdaderos, sin poder validarlos. Pero el proceder así, reduce drásticamente el número de parámetros a estimar.

Introducir una FR y considerar que los reclutamientos y biomasa desovante que los produjeron, se adecúan a dicha función puede generar una restricción importante. Una posibilidad es relajar este condicionamiento y perturbar aleatoriamente los valores estimados por la FR.

Se verá en este trabajo una posibilidad para hacer esto. También se analizan los supuestos subyacentes al usar funciones de reclutamiento, en los modelos de dinámica de poblaciones estructuradas por edades. El supuesto de distribución log-normal de errores se plantea formalmente y se discute sobre la manera de considerarlo en el modelo de dinámica poblacional. Se discute sobre los efectos de los supuestos no validables, que conlleva tanto el uso de funciones de reclutamiento, como de introducir la aleatoriedad en los desvíos de los reclutamientos de la misma y el impacto sobre las inferencias a largo plazo.

Se propone una alternativa metodológica de generación de reclutamientos con distribución de probabilidades log-normal y función media dada. Este trabajo va dirigido a todos los investigadores que trabajan con modelos de dinámica de poblaciones explotadas.

Métodos

Introducción de aleatoriedad en los modelos de reclutamiento

Algunos modelos de reclutamiento incluyen a las familias de funciones de Beverton y Holt y a la función de Ricker:

$R_{t+1} = \frac{\alpha\, SB_t}{\beta + SB_t}, \alpha, \beta > 0$ Beverton y Holt

y

$R_{t+1} = \alpha\, SB_t\, e^{-\beta\, SB_t}, \alpha, \beta > 0$ Ricker

y donde SB_t es la biomasa de hembras desovantes en el tiempo t, en peso, y R_{t+1} es el reclutamiento esperado en número de individuos, en el tiempo $t+1$.

Para procesos no compensatorios del reclutamiento, las FR suelen tener más de dos parámetros (por ejemplo: $R_{t+1} = \frac{\alpha\, SB_t^{\delta}}{\beta + SB_t^{\delta}}$, con $\alpha, \beta, \delta > 0$ es la función de Beverton y Holt generalizada).

Existen varias familias de funciones adicionales que pueden ser utilizadas, en casos particulares.

Algunos autores modelan la relación entre los reclutas y la biomasa desovante que les dio origen asumiendo normalidad, log-normalidad, gamma, e inclusive errores con distribución de Poisson. Pero el supuesto más común es que los errores o desvíos del reclutamiento con el reclutamiento medio, son variables aleatorias log-normales (Hilborn y Walters, 1992). Este supuesto comenzó a utilizarse con Allen (1973), basado en resultados empíricos. La función de reclutamiento oficia de valor medio (esperanza matemática) de la distribución de probabilidades.

Observar que cualquier supuesto de distribución de probabilidades de los errores o desvíos, será asumido seguramente sin validación, ya que es muy difícil que existan valores históricos de reclutamiento y suficientes de los mismos para validar este supuesto en cada tiempo. Además, los valores de reclutamiento son, en general, estimaciones del modelo de dinámica poblacional. Si existen campañas de investigación dirigidas al reclutamiento, en el mejor de los casos se podrá estimar un índice de abundancia de reclutas, pero no la abundancia.

Entonces, la asignación de una distribución de probabilidades como log-normal, también introduce un supuesto no validable, que puede generar problemas si no es válido.

La necesidad de establecer una FR y una distribución de probabilidades de los desvíos del reclutamiento de dicha función media, esto para algunos modelos de dinámica poblacional estructurada, genera así un aumento de la incertidumbre estructural.

De ser posible (cuando los datos lo permiten), es recomendable trabajar con modelos alternativos donde no sea necesario realizar supuestos no validables, por lo menos en cuestiones que afectan a la dinámica asintótica. Dentro del estudio de la dinámica asintótica está el análisis de los estados de equilibrio, y entonces el uso de supuestos no validables impactará sobre cualquier análisis de sostenibilidad biológica que se realice, generando mayor incertidumbre estructural (no medible).

Resultados

Generación de reclutamientos

El objetivo es generar una variable aleatoria que se denominará “reclutamiento” de manera que la esperanza matemática sea el valor dado por la función de reclutamiento y que esta variable aleatoria tenga distribución log-normal.

Sean:

R_t: reclutamiento medio, calculado a partir de una función de reclutamiento

r_t: una variable aleatoria normal de media cero y desvío estándar el del logaritmo del reclutamiento en el tiempo t.
σ_t: desvío estándar del logaritmo del reclutamiento en el tiempo t.
A esta variable aleatoria la llamaremos desvío del reclutamiento medio. $r_t \sim N(0, \sigma_t^2)$

Sea $R_t^0 = R_t e^{r_t}$. Este R_t^0 es una variable aleatoria log-normal.

$E(R_t^0) = R_t\ E(e^{r_t}) = R_t\ e^0 e^{\sigma_t^2/2} = R_t\ e^{\sigma_t^2/2}$ (por D) del Apéndice)

De esta forma, la variable aleatoria R_t^0 es sesgada. Se corrige la misma para generar reclutamientos insesgados respecto del valor medio dado por la función de reclutamiento.

Sea $R_t^* = R_t^0\ e^{-\sigma_t^2/2}$. Veamos que $E(R_t^*) = E\left(R_t^0\ e^{-\frac{\sigma_t^2}{2}}\right) = R_t$. Que es lo deseado.

$$E\left(R_t^0\ e^{-\frac{\sigma_t^2}{2}}\right) = E(R_t^0) e^{-\sigma_t^2/2} = R_t\ e^{\sigma_t^2/2}\ e^{-\sigma_t^2/2} = R_t$$

Por lo tanto, $R_t^* = R_t\ e^{r_t - \sigma_t^2/2}$, es una variable aleatoria log-normal, de media R_t el valor de reclutamiento medio para el tiempo t.♦

Discusión

Notar que el procedimiento dado permite incorporar aleatoriedad al reclutamiento medio calculado con la función de reclutamiento. Dicha aleatoriedad considera a los reclutamientos provenientes de variables aleatorias con distribución log-normal, con media y varianza dependiendo del tiempo t.

Para cada tiempo es necesario fijar el desvío estándar del logaritmo del reclutamiento. A priori, este desvío estándar no se conoce. Si se introdujeran como parámetros a estimar por el modelo, no habría ganancia de reducción de parámetros. Habrá que asumir algún supuesto para mantener la reducción significativa de parámetros. Por ejemplo, asumir que $\sigma_t = \sigma$ para todo tiempo es una opción. El valor de σ puede fijarse con algún sentido práctico. Un valor grande liberará más a los reclutamientos que un valor pequeño. También podría estimarse este σ como parámetro del modelo de dinámica poblacional.

Pero para dar plasticidad a las estimaciones del modelo de dinámica poblacional, es conveniente incorporar una simulación de Monte Carlo, que facilite la generación aleatoria de los reclutamientos y permita mejorar el proceso de optimización.

Un problema surge al asumir una determinada familia de funciones de reclutamiento y a su vez asumir los reclutamientos aleatorios con σ_t grande. En este caso puede generase mucha variabilidad de manera que sea muy dificultoso estimar los parámetros de la FR y del modelo de dinámica poblacional.

En el proceso de estimación de parámetros, se requeriría, cada vez que se evalúa la función de verosimilitud (enfoque frecuentista y bayesiano), con un vector de parámetros, generar aleatoriamente valores de r_t para cada t. Suficiente cantidad de valores como para facilitar una mejora en las estimaciones. Una posibilidad es generar aleatoriamente una cantidad determinada de valores y lograr quedarse con los valores de r_t que brinde un óptimo de la función objetivo (para el enfoque frecuentista) para dicha evaluación con el vector de parámetros. Y otra posibilidad es generar solo un valor de r_t para cada t. Ambas variantes pueden lograrse programando el modelo en alguna plataforma o lenguaje de programación. Pero notar que la simulación será parte de la definición del valor de la función objetivo (o de verosimilitud) en cada vector de parámetros.

En principio, se puede observar que la propuesta anterior contempla la generación de r_t para cada tiempo con distribución normal e independiente de otros tiempos. Sin embargo, debiera considerarse la posibilidad que estos desvíos no sean realmente independientes.

Podría ocurrir que la variabilidad en los reclutamientos dependa de factores abióticos similares que se dan en periodos de tiempo y luego cambian. Podría ocurrir entonces, periodos de menor variabilidad y periodos de mayor variabilidad. Considerar un valor constante de $\sigma_t = \sigma$, no sería conveniente en este caso.

Considerar la covariación para distintos tiempos de r_t, es difícil. Sólo con conocimiento extra de las variables que afectan al desvío del reclutamiento del valor medio, podría modelarse este efecto de covariación o por lo menos regular los valores de σ_t a priori (Subbey *et al*., 2014). Diferentes autores incorporan efectos del medio ambiente en los desvíos (ver, por ejemplo, Hilborn y Walters, 1992; Subbey *et al*., 2014). Pero no sólo variables ambientales pueden afectar la variabilidad del reclutamiento, sino que también efectos combinados del ambiente con el nivel de biomasa desovante y efectos de cambios en las fertilidades por gran explotación poblacional. También la estructura de edades de la biomasa de hembras desovante es muy importante para el éxito reproductivo (buen reclutamiento) (Aubone, 2015 a). Debe tenerse en cuenta que el planteo de una función de reclutamiento basada en relación de reclutas con la biomasa desovante que les dio origen, se basa en que la biomasa desovante es un indicador destacado de la producción de huevos y su supervivencia. Pero esto no siempre ocurre, y es importante considerar variables como la estructura de edades, el medio ambiente y la abundancia poblacional (Aubone, 2015 a), por ejemplo, como ya se discutió. Por otro lado, se asume cierta estabilidad en el ecosistema, para que la función de reclutamiento pueda existir.

Si se considerara a los valores de r_t, para los tiempos de evaluación, como parámetros del modelo, en el enfoque frecuentista, no hay realmente dependencia de estos valores estimados de la distribución normal, que se asume en la propuesta (no se los considera aleatorios). Además de asumir falsamente la aleatoriedad de cada desvío, este accionar no reduce el número de parámetros a estimar, con el posible problema de sobre-parametrización, como ya se dijo.

Con el enfoque bayesiano para la estimación de parámetros, la simulación será parte de toda la simulación numérica de un método SIR (*Sampling Importance Resampling Method*), por ejemplo. En este caso, incluso podrían tratarse los reclutamientos de manera totalmente libres y simularlos en el periodo de evaluación, sin condicionamientos. Se sugiere seguir este lineamiento.

Bibliografía

Allen, K. R. 1973. The influence of random fluctuations in the stock recruitment relationship on the economic return from salmon fisheries. Rapports et Proces-Verbaux des Reunions, Conseil International Pour L'Exploration de la Mer, 164: 350–359.

Aubone, A., 2015 a. Análisis ergódico de sostenibilidad biológica para poblaciones de peces: estructura de edades. 220 p. E-book (pdf), 16/06/2015. ISBN 978-987-33-7884-3; http://hdl.handle.net/1834/6885

Aubone, A., 2015 b. Dinámica de poblaciones de peces estructuradas. 226 p. E-book (pdf), 15/07/2015. ISBN 978-987-33-8150-8

Aubone, A. 2021. Notas de Matemática Pesquera. 460 p. E-book (pdf). (*En revisión final*)

Hilborn, R., and Walters, C. J. 1992. Quantitative fisheries stock assessment. Choice, dynamics and uncertainty. Chapman and Hall.

Subbey, S.; Devine, J.A.; Schaarschmidt, U. and Nash, R.D.M. 2014. Modelling and forecasting stock–recruitment: current and future perspectives. ICES Journal of Marine Science (2014), 71(8), 2307–2322. doi:10.1093/icesjms/fsu148

Walters, C. J., and Hilborn, R. 1976. Adaptive control of fishing systems. Journal of the Fisheries Board of Canada, 33: 145–159.

Apéndice

DISTRIBUCION LOGNORMAL

Sea X es una variable aleatoria cuyo logaritmo Y es normalmente distribuido con media η y varianza σ^2; esto es $y = \ln(x)$ y

$$p(y) = \frac{1}{\sqrt{2\pi\sigma}} e^{-(y-\eta)^2/(2\sigma^2)} \quad -\infty < \mathrm{y} < \infty$$

es la función de densidad de Y, entonces X se denomina una variable aleatoria *lognormal*.

Se define $\xi = e^{\eta}$.

<u>Ejercicio:</u>

A) encontrar la función de densidad de X
B) probar que la mediana de la distribución de X es ξ
C) probar que el modo de la distribución de X es $\xi\, e^{-\sigma^2}$
D) probar que $E(X) = \xi\, e^{\sigma^2/2}$
E) probar que $V(X) = e^{\sigma^2+2\eta}\,(e^{\sigma^2} - 1)$
F) probar que $V(Y) = \ln(1 + \frac{V(X)}{E(X)^2})$
G) probar que $E(Y) = \ln(\frac{E(X)}{\sqrt{1 + \frac{V(X)}{E(X)^2}}})$