FCEA FACULTAD DE CIENCIAS ECONÓMICAS Y DE ADMINISTRACIÓN

UNIVERSIDAD DE LA REPÚBLICA URUGUAY

Curso Demografía - Licenciatura en Estadística

Docentes:
Daniel Ciganda
Facundo Morini

13ra Clase
21 de Octubre de 2025

Simulación de Procesos Demográficos

En la siguiente unidad temática vamos a trabajar sobre algunas técnicas para simular procesos demográficos y modelar los resultados de esas simulaciones. Nos vamos a enfocar en los eventos primer hijo y fallecimiento.

El primer objetivo es comprender cómo podemos generar tiempos de espera individuales a cada uno de estos eventos a partir de unas funciones de riesgo acumulado.

Este ejercicio nos permitirá: introducir algunas ideas iniciales del análisis de supervivencia; entender mejor las tasas demográficas, su relación con la función de riesgo y la relación entre los niveles micro y macro de análisis.

Variable Aleatoria Discreta

Variable Aleatoria Discreta X Valores potenciales: $x_1, ..., x_k$

Función de Probabilidad:

$f(x_k) = P(X = x_k) \qquad k = 1, ..., K$

$$\sum_{k=1}^{K} f_k = \sum_{k=1}^{K} P(X = x_k) = 1$$

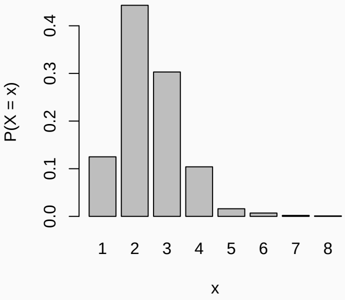
Función de Distribución:

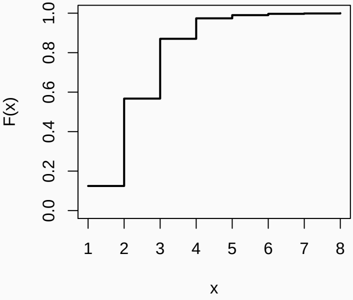
$$F(x) = P(X \leq x) = \sum_{x_k \leq x} f_k$$

Variable Aleatoria Discreta - Ejemplo

X: Nr. ideal de hijos (IVS: Paises Europeos 1990)

x_k	1	2	3	4	5	6	7	8
f_k	0.125	0.443	0.303	0.104	0.016	0.007	0.002	0.001



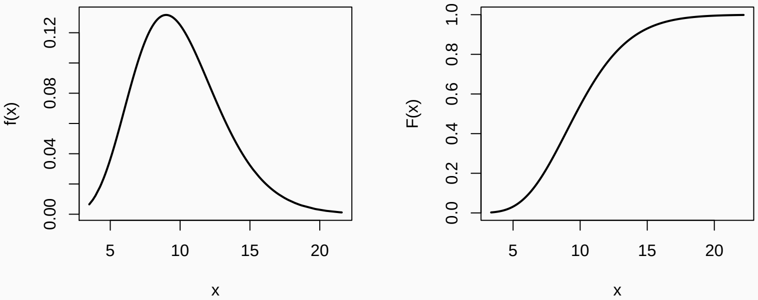


Función de Densidad: $f(x)$

$$P(a \leq X \leq b) = \int_a^b f(x)\, dx$$

Función de Distribución:

$$F(x) = P(X \leq x)$$



Funciones para simular en R

`runif`: Distribución Uniforme
`rnorm`: Distribución Normal
`rgamma`: Distribución Gamma
`rexp`: Distribución exponencial

Ejemplos:

```
runif(n=100, min = 0, max = 1)
rnorm(n=50, mean = 0, sd = 1)
```

Uno de los métodos para simular números aleatorios de una distribución:

Método de Inversión

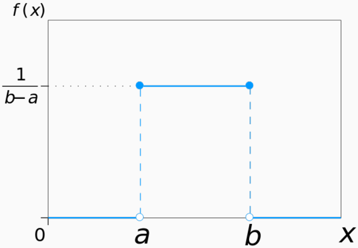


Figure 1: Distribución Uniforme

Método de Inversión

El objetivo es generar realizaciones de X.

Para ello, utilizamos la **función inversa** de $F(x)$, denotada como F^{-1}

- La función de distribución acumulada $F(x)$ mapea valores de x a probabilidades y en el intervalo $[0, 1]$.
- Su inversa $F^{-1}(y)$ nos permite ir desde una probabilidad y de regreso al valor x correspondiente.

$$F^{-1}(F(x)) = x \qquad \ln(e^x) = x$$

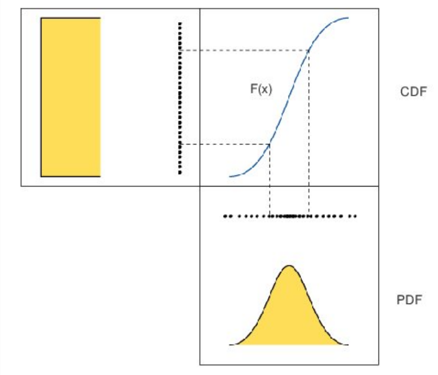


Figure 2: Método de Inversión

Método de Inversión

Para generar realizaciones de X:

1. Generamos valores aleatorios $u_1, \ldots, u_n \sim U(0,1)$.
2. Aplicamos la transformación: $x_i = F^{-1}(u_i) \quad \text{para } i = 1, \ldots, n$.

De esta manera, los x_i obtenidos siguen la distribución de probabilidad deseada.

Ejemplo:

Supongamos que queremos generar una variable aleatoria exponencial con parámetro λ:

- La función de distribución acumulada es $F(x) = 1 - e^{-\lambda x}$
- Su inversa es $F^{-1}(u) = -\frac{1}{\lambda}\ln(1-u)$

Entonces:

1. Generamos $u \sim U(0,1)$.
2. Calculamos $x = F^{-1}(u) = -\frac{1}{\lambda}\ln(1-u)$.

Generación de números aleatorios de la distribución exponencial a partir del método de la transformación inversa.

El método de inversión puede aplicarse tanto sobre la Función de Distribución como sobre la Función de Riesgo Acumulado cuando trabajamos con modelos de tiempo al evento.

Esto es relevante ya que las tasas específicas por edad de los eventos demográficos nos proporcionan la distribución de riesgo empírica bajo un supuesto de riesgo constante a intervalos.

Es decir, que podemos partir de unas tasas específicas de mortalidad para una cohorte, por ejemplo, y simular tiempos hasta la muerte usando el método de la transformación inversa.

Modelos de tiempo al evento

Análisis de sobrevida o análisis de historia de eventos

El tiempo al evento es entendido como una variable aleatoria caracterizada por una series de distribuciones:

- Función de distribución
- Densidad de probabilidad
- Función de supervivencia
- Función de riesgo

Realizaciones de la variable aleatoria "duración de la vida"

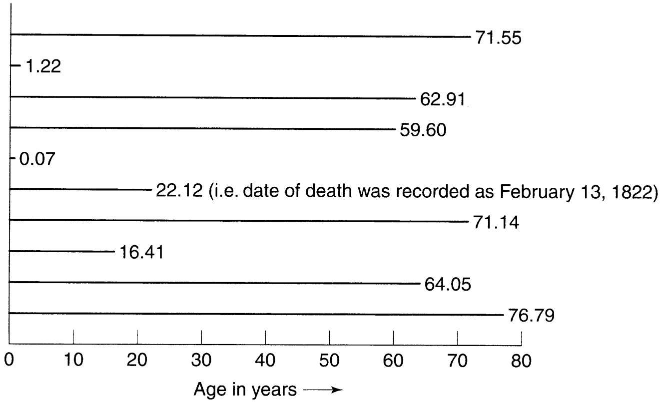


Figure 3.1 Age at death and life-lines of a hypothetical cohort of births (10 in all); date of birth: January 1, 1800

Modelos de Tiempo al Evento

Variable Aleatoria $T > 0$, continua.

Función de Distribución:

$$F(t) = P(T \leq t)$$

Probabilidad de que el evento ocurra antes que t.

Función de Supervivencia:

$$S(t) = P(T > t) = 1 - F(t)$$

Probabilidad de que el evento no haya ocurrido hasta el tiempo t.

Función de Densidad:

$$f(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t)}{\Delta t}$$

Modelos de Tiempo al Evento (continuación)

Función de Riesgo (Hazard):

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t} = \frac{f(t)}{S(t)}$$

Es la tasa instantánea de ocurrencia del evento en el tiempo t, dado que no ha ocurrido antes de t.

Función de Riesgo Acumulado:

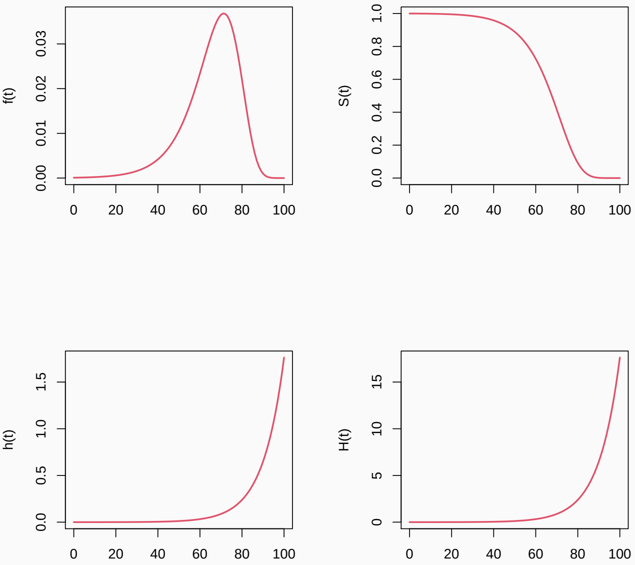
$$H(t) = \int_0^t h(s)\, ds$$

Representa la acumulación del riesgo hasta el tiempo t.

Además, podemos expresar la función de supervivencia en términos del riesgo acumulado:

$$S(t) = \exp\{-H(t)\}$$

Ej. Gompertz


f(t)
0.00
0.01
0.02
0.03
0
20
40
60
80
100
S(t)
0.0
0.2
0.4
0.6
0.8
1.0
0
20
40
60
80
100
h(t)
0.0
0.5
1.0
1.5
0
20
40
60
80
100
H(t)
0
5
10
15
0
20
40
60
80
100

Estimador de Máxima Verosimilitud de la Función de Riesgo

Cuando asumimos que el riesgo es constante en cada intervalo de edad, el vector de tasas de ocurrencia/exposición por edad observadas representa la función de riesgo que caracteriza a la variable aleatoria *tiempo al evento*.

Esto es así ya que, bajo el supuesto de riesgo constante, el Estimador de Máxima Verosimilitud (EMV) del riesgo de un evento es la cantidad de eventos observados sobre el tiempo de exposición al riesgo del evento, es decir la tasa de ocurrencia/exposición.

$$m_x = \frac{D_x}{L_x}$$

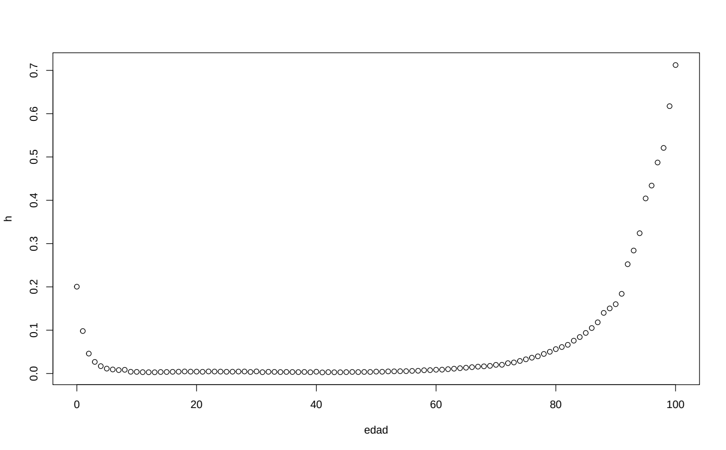


Figure 3: Tasas de Mortalidad por Edad de una Cohorte

$$f_{i,x} = \frac{B_{i,x}}{L_{i-1,x}}$$

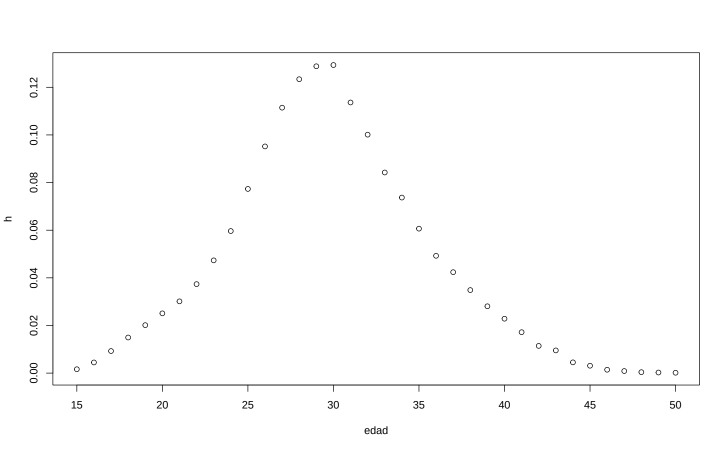


Figure 4: Tasas Condicionales de Fecundidad por Edad de una Cohorte - Primer Hijo

Modelo a intervalos riesgo constante (Piecewise-constant hazard)

Definimos los siguientes intervalos:

$$[\tau_0 = 0, \tau_1)[\tau_1, \tau_2), ..., [\tau_{p-1}, \tau_p), [\tau_p, \infty)$$

En cada intervalo $I_j, j = 1, ..., p+1$ asumimos que el riesgo es constante: λ_j

$$h(t) = \sum_{j=1}^{p+1} \lambda_j \, \mathbb{1}_{[\tau_{j-1}, \tau_j)}(t)$$

Laboratorio en R

Cálculo de la función de riesgo, riesgo acumulado y función de supervivencia

Script *“survival.R”*