



FACULTAD DE  
CIENCIAS ECONÓMICAS  
Y DE ADMINISTRACIÓN



UNIVERSIDAD  
DE LA REPÚBLICA  
URUGUAY

## Microsimulación

---

Docente: Daniel Ciganda

8<sup>va</sup> Clase

22 de Octubre de 2025

- En la primera parte, trabajamos con **Modelos Basados en Agentes (MBA/ABM)**.
  - Enfatizan **interacciones explícitas** y *dinámica emergente*.
  - Útiles para estudiar mecanismos, redes y procesos fuera de equilibrio.
- La **microsimulación tradicional (MS)** también es *bottom-up*, pero:
  - Se centra en **proyectar trayectorias de vida** bajo **reglas/leyes** y **transiciones estimadas**.
  - Interacciones entre unidades suelen ser *indirectas* (vía totales, mercados, cupos).
- **Diferencia clave:** ABM pregunta “¿qué *emerge* de la interacción?”; MS pregunta “¿*cómo* evolucionará la población si operan estas **transiciones y reglas**?”.

# ¿Qué es la Microsimulación?

## Definición

Técnica que simula **unidades individuales** (personas, hogares, empresas) a lo largo del **tiempo** aplicando **probabilidades de transición** y **reglas** para producir resultados **agregados y distributivos**.

## Idea central:

- Partimos de una **población base** (muestra representativa o población sintética).
- Para cada unidad y periodo, aplicamos **transiciones** y **actualizamos** su vector de estado.
- Al **agregar** miles/millones de trayectorias, obtenemos indicadores macro y *distribuciones* (no solo promedios).

- **Años 50–70:** Guy **Orcutt** propone modelar “desde abajo”: individuos/hogares en lugar de promedios agregados.
- **Motivación:** las políticas (impuestos, pensiones, elegibilidad) tienen **efectos no lineales** y **heterogéneos**.
  - Los promedios ocultan **umbrales** y **cliffs** (p. ej. pérdida de beneficios).
- **Propuesta:** simular **micro-decisiones y eventos** y luego **agregar**.
  - Permite medir **impacto distributivo**, **trayectorias** y **casuísticas** con más realismo.

# Los Componentes Clave de un Modelo de MS

Un modelo de microsimulación tradicional tiene tres ingredientes principales:

## 1. Población Base Microdatos

representativos (censo, encuesta, registros) o **población sintética**.

Incluye variables demográficas y socioeconómicas.

## 2. Módulos de Transición

**Ecuaciones** (ej. logit, hazard) que dan  $P(\text{evento} \mid X)$ .

Ej.:

$$\text{logit}(P(\text{muerte})) = \beta_0 + \beta_1 \text{edad} + \beta_2 \text{sexo} + \dots$$

## 3. Motor de Simulación

Itera en el **tiempo**, simulando eventos via Monte Carlo, **actualiza estados** y computa variables derivadas (impuestos, pensión, elegibilidad).

- **Estática vs. Dinámica**

- *Estática*: un periodo; reponderación y reglas  $\Rightarrow$  impactos distributivos inmediatos.
- *Dinámica*: múltiples periodos; **memoria** y **trayectorias** (curso de vida).

- **Tiempo**: *discreto* (anual/mensual) vs. *dirigido por eventos* (tiempo continuo).
- **Población**: *cerrada* (sin nac./mig.) vs. *abierta* (entradas/salidas).

- **Discreto ( $t=1,2,\dots$ ):** más simple; orden de eventos importa.
  - Secuencias típicas: *edad*  $\rightarrow$  *mortalidad*  $\rightarrow$  *estado civil*  $\rightarrow$  *fecundidad*  $\rightarrow$  *empleo*  $\rightarrow$  *ingresos*  $\rightarrow$  *impuestos/beneficios*.
- **Por eventos (tiempo continuo):** usa *tasas de riesgo* y *competencia de riesgos*.
  - $h_k(t \mid X) = \lim_{\Delta t \rightarrow 0} \frac{P(\text{evento } k \text{ en } [t, t+\Delta t] \mid \text{no evento antes}, X)}{\Delta t}$
- **Resolución de conflictos:** eventos simultáneos, prioridades, reglas de desempate.

## El Loop de Microsimulación (Workflow)

1. **Inicialización:** limpiar/armonizar microdatos; crear población base; fijar semillas.
2. **Cómputo de transiciones:** para cada unidad,  $P(\text{evento} \mid X_t)$  según módulos.
3. **Sorteos Monte Carlo:** realizar draws; manejar *riesgos competitivos*.
4. **Actualización:** cambiar estados; recalcular variables derivadas (impuestos, elegibilidad, pensión).
5. **Alineación/Calibración:** ajustar a totales/targets (población, empleo, mortalidad, recaudación).
6. **Escenarios:** comparar *baseline* vs. *reformas/políticas*.

- **Fuentes:** encuestas (ingresos, empleo, salud), censos, registros administrativos, historias clínicas/otros datasets específicos.
- **Construcción de base:**
  - *Uso directo* de microdatos (limpieza, imputaciones, armonización).
  - *Reponderación* a márgenes conocidos por edad/sexo/región, etc.
- **Parámetros:** estimación de regresiones/hazards, matrices de transición, o reglas normativas.

- **Tributaria/Beneficios (estática):** impuestos, transferencias, pobreza, incentivos al trabajo.
  - Cambios en tramos, créditos, elegibilidad; *incidencia* por decil, hogares con niños, etc.
- **Demografía/Envejecimiento (dinámica):** uniones, fecundidad, educación, migración, mortalidad.
  - Proyecciones de estructura por edad, hogares, demanda de servicios.
- **Salud (historia natural):** costo-efectividad tratamientos, carga de enfermedad.
- **Mercado laboral y pensiones:** carreras, aportes, tasas de reemplazo, edad de retiro.

## Fortalezas

- Detalle **distributivo** y **heterogeneidad**.
- Reglas **realistas** (umbrales, elegibilidad, topes).
- **Contrafactuales** limpios para evaluación de políticas.
- Integra **encuestas** y **registros** con alineación.

## Limitaciones

- **Intensiva en datos** y supuestos de transición.
- **Sensibilidad** a alineación/parametrización.
- Costos de **cómputo** y **validación** multicapas.
- Interacciones explícitas **limitadas** (vs. ABM).

- **Micro-validación:** distribución de variables clave, correlaciones, coherencia interna.
- **Macro-validación:** totales y tasas vs. series externas (backcasting).
- **Pruebas de sensibilidad:** parámetros, orden de eventos, supuestos de alineación.
- **Reproducibilidad:** *scripts* orquestados, *logs*, y reportes automatizados.

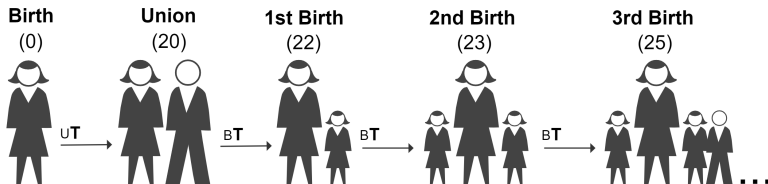
## ¿Cuándo usar (y cuándo no) Microsimulación?

- **Usar MS cuando:**

- Importa la **distribución** (no solo el promedio).
- Las reglas tienen **umbrales/no linealidades**.
- Existe **microdato** razonable y **targets** agregados confiables.

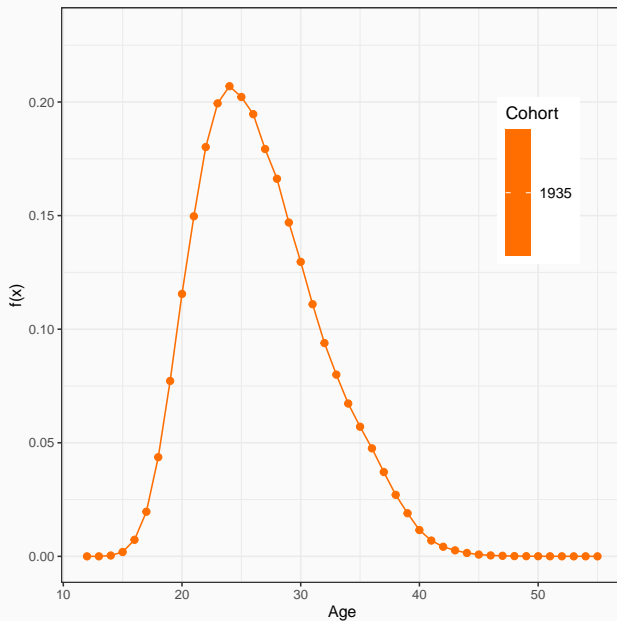
- **Evitar MS si:**

- La pregunta es puramente **agregada** y sin heterogeneidad relevante.
- No hay datos para **parametrizar** ni **validar**.
- Se requieren **interacciones complejas** (considerar ABM).

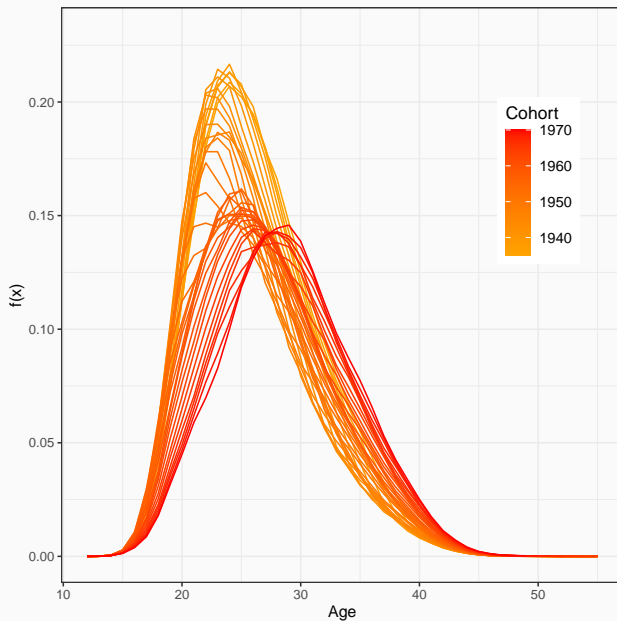


**Figure 1:** El Proceso Reproductivo en un Contexto de Fecundidad Natural

**Figure 2:** Tasas Específicas de Fecundidad por Edad de una Cohorte, 1935, Francia



**Figure 3:** Tasas Esp. de Fecundidad por Edad de una Cohorte, 1935-1970, Francia



Tasas específicas de fecundidad por edad:

$${}_n f_x = \frac{{}_n B_x}{{}_n E_x}$$

$n$  = el tamaño del intervalo

$B$  = nacimientos asociados a mujeres con edad en el intervalo  $[x, x + n)$

$E$  = nro. de años persona aportados en el intervalo  $[x, x + n)$

Tasa Global de Fecundidad:

$$TGF = n \cdot \sum_{x=\alpha}^{\beta-n} {}_n f_x$$

# Modelos de Tasas Específicas de Fecundidad (ASFR): Enfoques

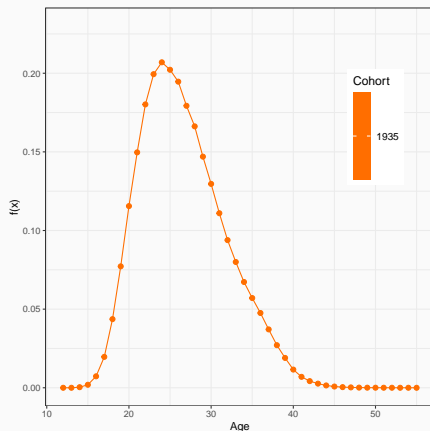
El problema es generalmente abordado desde el nivel **agregado**.

## Enfoques

- No Paramétricos — Splines
- Paramétricos — Coale and Trussell (1978)

Modelos evaluados con respecto a:

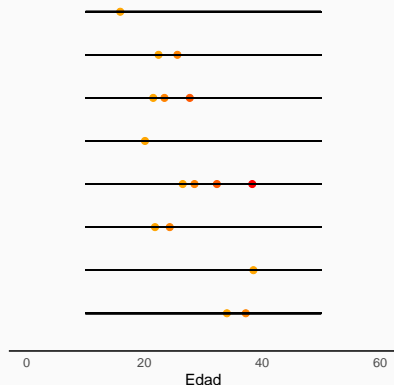
- \* Qué tan bien se ajustan a los datos
- \* La **interpretabilidad** de sus parámetros



# El Modelo: El Proceso Reproductivo en Régimen de Fecundidad Natural

4 piezas fundamentales:

1. Inicio del Proceso: **Edad al matrimonio**
2. Probabilidad de concebir: **Fecundabilidad**, Gini (1924)
3. Período de **no-susceptibilidad**: embarazo + amenorrea postparto
4. **Fin del Proceso**: Disminución del riesgo de concebir hasta la menopausia

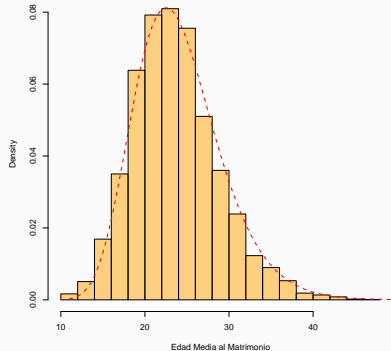


**Figure 4:** Trayectorias Reproductivas

# El Modelo: El Proceso Reproductivo en Régimen de Fecundidad Natural

## 1. Inicio del Proceso: Edad al matrimonio

$$X_m \sim \text{Lognormal}(\mu_m, \sigma_m)$$



2. **Fecundabilidad:** *“La probabilidad de concebir de una mujer casada en un mes, en ausencia de prácticas Malthusianas.”*

Llamamos  $\phi$  a la probabilidad de concebir en el mes.



**Figure 5:** Corrado Gini

# Modelo del Primer Nacimiento

Datos: Frecuencia mensual de nacimientos después del casamiento

Si denotamos fecundabilidad como  $\phi$  y asumimos que las concepciones empiezan inmediatamente después del casamiento:

Proporción de concepciones 1<sup>er</sup> mes =  $\phi \implies \phi$  nacimientos 9 meses después

Una proporción de mujeres  $1 - \phi$  no concebirá y lo intentará el mes siguiente

Durante el 2<sup>do</sup> mes una proporción  $(1 - \phi)\phi$  concebirá y dará a luz 9 meses más tarde

Proporción de concepciones en el mes  $n$  después de la unión será:

$$(1 - \phi)^n \phi$$

Para:  $n = 0, 1, 2, \dots$

# Distribución Geométrica

Distribución de probabilidad discreta:

- **Primera versión:**  $X$  representa el número de ensayos Bernoulli necesarios para obtener el primer éxito:

$$P(X = k) = (1 - p)^{k-1} p$$

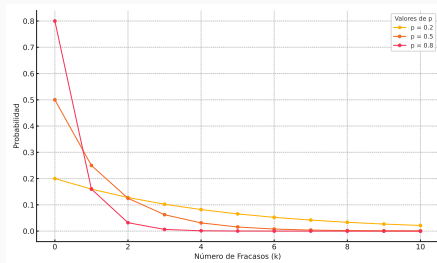
Para:  $k = 1, 2, 3, \dots$

- **Segunda versión:**  $Y = X - 1$  representa el número de fracasos antes del primer éxito.

$$P(Y = k) = (1 - p)^k p$$

Para:  $k = 0, 1, 2, \dots$

**Figure 6:** Distribución Geométrica



Es un caso especial de la distribución Binomial Negativa que modela la probabilidad de  $n$  éxitos.

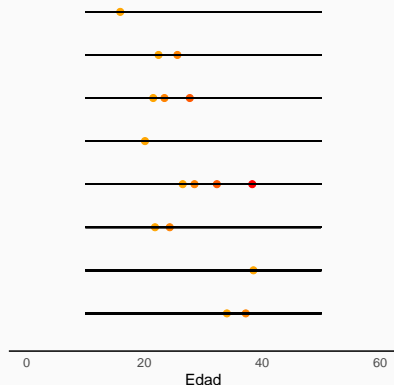
En R `rbinom()`

Simulación de tiempos de espera y edades a distintas paridades.

# El Modelo: El Proceso Reproductivo en Régimen de Fecundidad Natural

4 piezas fundamentales:

1. Inicio del Proceso: **Edad al matrimonio**
2. Probabilidad de concebir: **Fecundabilidad**, Gini (1924)
3. Período de **no-susceptibilidad**: embarazo + amenorrea postparto
4. **Fin del Proceso**: Disminución del riesgo de concebir hasta la menopausia

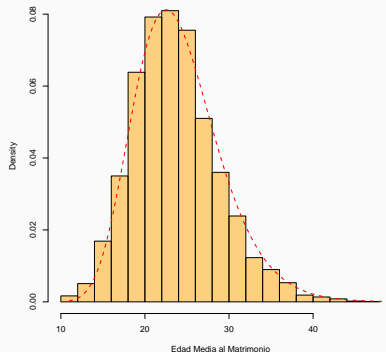


**Figure 7:** Trayectorias Reproductivas

# El Modelo: El Proceso Reproductivo en Régimen de Fecundidad Natural

## 1. Inicio del Proceso: Edad al matrimonio

$$X_m \sim \text{Lognormal}(\mu_m, \sigma_m)$$



2. **Fecundabilidad:** *“La probabilidad de concebir de una mujer casada en un mes, en ausencia de prácticas Malthusianas.”*

Llamamos  $\phi$  a la probabilidad de concebir en el mes. Tenemos una concepción cuando:

$$\phi \geq u$$

con  $u \sim \mathcal{U}(0, 1)$



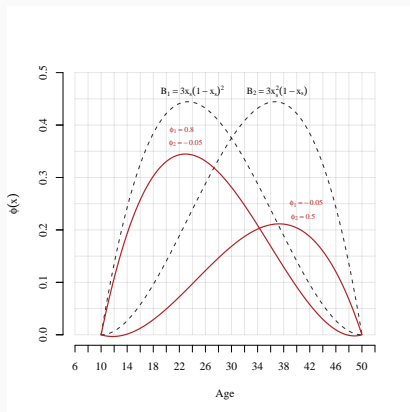
**Figure 8:** Corrado Gini

## 4. Estructura por edad del riesgo de concebir

Para modelar la evolución de la probabilidad de concebir con la edad usamos dos funciones base y dos coeficientes.

Los coeficientes  $\beta$  permiten manipular la **forma** de la curva asignando un “peso” a cada una de las funciones base.

$$\phi(t) = B_1^3(t)\phi_1 + B_2^3(t)\phi_2$$



## Parámetros:

1. Media edad al matrimonio  $\mu_m$
2. Desv. St. edad al matrimonio  $\sigma_m$
3. Máximo fecundabilidad  $\phi_1$
4. Caída fecundabilidad  $\phi_2$
5. Período de no suceptibilidad  $\delta$

### initialization

create initial cohort of women

assign  $x_i$ ,  $\phi(x_i)$ , and  $wts_i$

**while**  $time < end\_time$  **do**

*update*  $t$

*update*  $x_i$  and  $wts_i$

*update*  $\phi(x_i)$

**married at**  $t$

        set marriage status = 1

        set age at union = age

**back to susceptible state at**  $t$

        set susceptible status = 1

**conception**

        draw  $u_i \sim \text{Uniform}(0, 1)$

**if**  $u_i < \phi(x_i)$  **then**

        set  $wt$  to birth = pregnancy (9  
            months)

        set susceptible status = 0

        set  $wt$  to susceptible =  
            pregnancy + length of  
            non-susceptibility period

**end**

**birth at**  $t$

        children = children + 1

        age at  $j^{th}$  birth = age

**end**

compute  $f^{\dagger}(x)$  from reproductive histories

### Se parecen a MS (orientados por los datos)

- Diseñados para **ajustar datos observados** (*targets* agregados: ASFR, TFR, edad al matrimonio).
- **Menor necesidad de microdatos** ricos: se calibran con **tablas/series publicadas**.
- **Sin** redes, **sin** retroalimentaciones complejas, **sin** interacciones explícitas.

### Se parecen a ABM (mecanicistas)

- **Generativos**: especifican mecanismos simples (fecundabilidad, no-susceptibilidad, edad al matrimonio).
- Modelado de mecanismos en lugar de modelado estadístico de probabilidades de transición.

### Idea clave

Son modelos **intermedios**: **orientados a datos** y **mecanicistas**, estimados con **agregados**, útiles cuando no disponemos de microdatos detallados pero queremos trayectorias coherentes con la evidencia. **Parámetros** con significado demográfico directo).

## Simular un evento con probabilidad $\phi$

- Queremos decidir si ocurre un evento con probabilidad  $\phi \in [0, 1]$  (p.ej., concepción en un mes).
- Procedimiento (**Bernoulli por uniforme**):
  1. Sortear  $u \sim \mathcal{U}(0, 1)$ .
  2. Si  $u < \phi \Rightarrow$  **evento**; si  $u \geq \phi \Rightarrow$  **no evento**.
- **Intuición:** la probabilidad de caer en  $[0, \phi]$  dentro de  $[0, 1]$  es exactamente  $\phi$ .



## ¿Por qué funciona? Frecuencia de eventos $\approx \phi$

- Si repetimos el sorteo muchas veces, la proporción de  $u < \phi$  se acerca a  $\phi$ .
- Debajo:  $N$  sorteos de  $u \sim \mathcal{U}(0, 1)$ . Puntos verdes:  $u < \phi$ ; grises:  $u \geq \phi$ .

$N = 500$  sorteos, eventos = 173  $\hat{p} = 0.34999 \approx 0.35$ .

