

VALUTAZIONI ATTUARIALI MEDIANTE SIMULAZIONE STOCASTICA

Il ricorso alla simulazione stocastica può risultare utile quando è molto complesso o impossibile ottenere valutazioni in forma chiusa.

Ad esempio, consideriamo il seguente problema, siano dati X con funzione di ripartizione F_X e una funzione φ , poniamo $Y = \varphi(X)$. Può essere molto complesso ottenere valutazioni per Y . Analogamente, per $Y = \varphi(X_1, \dots, X_n)$, con assegnata $F_{X_1, \dots, X_n}$, o per $Y = \varphi(X_1, X_2, \dots)$, con $\{X_1, X_2, \dots\}$ processo di legge assegnata.

Un caso di interesse, sia

$$X = \sum_{i=1}^N Y_i = \varphi(N, Y_1, Y_2, \dots),$$

con **distribuzione composta** assegnata. Formalmente, si ha

$$F_X(x) = \sum_{n=0}^{+\infty} Pr(N = n) F_Y^{*(n)}(x),$$

ma operativamente F_X è ottenibile solo in casi molto particolari.

La **simulazione** di un ente aleatorio, di legge assegnata, consiste nella “sostituzione” di tale ente con un altro che riproduca **lo stesso stato di incertezza** dell’ente originale ovvero con un ente al quale si ritenga di assegnare **la medesima valutazione probabilistica dell’ente originale**:

- $X \sim F_X$ $\rightarrow X'$ tale che $X' =^d X$, $F_{X'} = F_X$,
- $\mathbf{X} = (X_1, \dots, X_n) \sim F_X$ $\rightarrow \mathbf{X}' = (X'_1, \dots, X'_n)$ tale che $\mathbf{X}' =^d \mathbf{X}$, $F_{\mathbf{X}'} = F_X$,
- $\{X_1, X_2, \dots\} \sim \mathcal{F}$ $\rightarrow \{X'_1, X'_2, \dots\}$ tale che $\{X'_1, X'_2, \dots\} =^d \{X_1, X_2, \dots\}$, $\mathcal{F}' = \mathcal{F}$.

Se X' è osservabile e x' è la determinazione, il valore osservato, di X' , x' è **una determinazione simulata** di X . Analogamente negli altri casi.

Se $Y = \varphi(X_1, \dots, X_n)$ ($Y = \varphi(X)$, $Y = \varphi(X_1, X_2, \dots)$), con assegnata $F_{X_1, \dots, X_n}$, per simulare Y si può simulare $(X_1, \dots, X_n)$, mediante $(X'_1, \dots, X'_n) =^d (X_1, \dots, X_n)$, e porre $Y' = \varphi(X'_1, \dots, X'_n)$. Essendo $Y' =^d Y$, Y' simula Y .

Se $(X'_1, \dots, X'_n)$ è osservabile e $(x'_1, \dots, x'_n)$ la determinazione di $(X'_1, \dots, X'_n)$, $y' = \varphi(x'_1, \dots, x'_n)$ è una determinazione simulata di Y .

Mediante più determinazioni simulate si possono ottenere stime dei momenti della distribuzione e della funzione di ripartizione di $Y = \varphi(X_1, \dots, X_n)$: se $\mathbf{X}'_1, \dots, \mathbf{X}'_m$ simulano $\mathbf{X} = (X_1, \dots, X_n)$ e sono stocasticamente indipendenti $\Rightarrow$

$$Y'_i = \varphi(\mathbf{X}'_i), \quad i = 1, \dots, m,$$

simulano Y e sono stocasticamente indipendenti, sono un campione casuale della distribuzione di Y . Allora, la media campionaria e la varianza campionaria,

$$\bar{Y}'_m = \frac{1}{m} \sum_{i=1}^m Y'_i, \quad S_m^2 = \frac{1}{m-1} \sum_{i=1}^m (Y'_i - \bar{Y}'_m)^2,$$

sono stimatori non distorti e consistenti del valore atteso e della varianza. La funzione di ripartizione empirica,

$$\tilde{F}_m(y) = \frac{1}{m} \sum_{i=1}^m |Y'_i \leq y|,$$

è stimatore della funzione di ripartizione.

Se $X'_1, \dots, X'_m$ sono osservabili e $x'_1, \dots, x'_m$ sono i valori osservati, allora $y'_i = \varphi(x'_i)$, $i = 1, \dots, m$, è il campione osservato della distribuzione di Y . Allora, se m è “sufficientemente” elevato,

- $\frac{1}{m} \sum_{i=1}^m y'_i$ è una stima di $E(Y)$,
- $\frac{1}{m-1} \sum_{i=1}^m (y'_i - \bar{y}'_m)^2$, con $\bar{y}'_m = \frac{1}{m} \sum_{i=1}^m y'_i$, è una stima di $var(Y)$,
- $F_m(y) = \frac{\text{card}\{i \mid y'_i \leq y\}}{m}$, $y \in \mathbb{R}$, è una stima della funzione di ripartizione di Y .

Per ottenere valutazioni probabilistiche di un ente aleatorio (numero o vettore aleatorio, processo stocastico) mediante simulazione stocastica si usano valori osservati di enti che simulano l'originale.

Affinché la simulazione sia efficace, è necessario che gli enti che simulano l'ente originale siano in numero elevato, facilmente e rapidamente osservabili.

Simulazione di un numero aleatorio

Sia X con funzione di ripartizione F_X assegnata.

Se U con determinazioni possibili in $]0,1[$ è tale che $U \sim Unif(0,1)$, allora, $X' = F_X^{\leftarrow}(U)$ è tale che $X' \stackrel{d}{=} X$. Pertanto, $X' = F_X^{\leftarrow}(U)$ simula X .

1) Se X è discreto con determinazioni possibili $x_1 < x_2 < \dots$ e $Pr(X = x_i) = p_i$, si ha $F_X(x) = \sum_{i: x_i \leq x} p_i$ e $X' = F_X^{\leftarrow}(U)$ è tale che

$$X' = x_i \Leftrightarrow F_X(x_{i-1}) < U \leq F_X(x_i).$$

Se U è osservabile e u la sua determinazione con $F_X(x_{i-1}) < u \leq F_X(x_i)$, la corrispondente determinazione simulata di X è $x' = F_X^{\leftarrow}(u) = x_i$.

2) Se F_X è strettamente crescente e continua dove è diversa da 0 e 1, allora esiste F_X^{-1} definita in $]0,1[$ e $X' = F_X^{-1}(U)$ simula X .

Se U è osservabile e u la sua determinazione, la corrispondente determinazione simulata di X è $x' = F_X^{-1}(u)$.

3) In generale, $X' = F_X^{\leftarrow}(U)$ simula X .

Se U è osservabile e u la sua determinazione, la corrispondente determinazione simulata di X è $x' = F_X^{\leftarrow}(u) = \inf\{x \in \mathbb{R} \mid F_X(x) \geq u\}$.

Dunque, la simulazione di un numero aleatorio X si può realizzare tramite U con determinazioni in $]0,1[$, con $U \sim Unif(0,1)$.

Per ottenere m numeri aleatori $X'_1, \dots, X'_m$ iid che simulano X , si può ricorrere ad una sequenza $U_1, \dots, U_m$ iid, con $U_i \sim Unif(0,1)$, e porre $X'_1 = F_X^{\leftarrow}(U_1), \dots, X'_m = F_X^{\leftarrow}(U_m)$. Affinché la simulazione sia efficace, è necessario che m sia elevato e che $U_1, \dots, U_m$ siano facilmente e rapidamente osservabili.

Numeri pseudo-casuali. I numeri pseudo-casuali sono numeri generati da un **algoritmo deterministico** che produce una sequenza di numeri che soddisfano opportuni test statistici, tali da poter essere visti come determinazioni di un campione $U_1, \dots, U_m$ iid, con $U_i \sim Unif(0,1)$. Ci sono diversi generatori di numeri pseudo-casuali che differiscono per il tipo di algoritmo usato. Un generatore deve essere inizializzato assegnando un **seme**. Ogni volta che si usa lo stesso seme si ottiene la stessa sequenza di numeri.

Notiamo che,

- nel caso 1), il problema di dover ricercare l'intervallo $]F_X(x_{i-1}), F_X(x_i)]$ in cui cade u può richiedere molto tempo se si necessita di molte determinazioni simulate,
- nel caso 2), bisogna che sia determinabile la funzione F_X^{-1} , ciò in molti casi non è possibile, ad esempio se $X \sim N(\mu, \sigma^2)$, $X \sim \text{gamma}(\alpha, \rho)$,
- nel caso 3) sono presenti entrambi i precedenti problemi.

In generale, si deve ricorrere ad opportuni algoritmi.

Simulazione di un vettore aleatorio o di un processo stocastico a parametro discreto

Sia $(X_1, \dots, X_n)$ con funzione di ripartizione congiunta $F_{X_1, \dots, X_n}$ o $\{X_1, X_2, \dots\}$ con legge $\mathcal{F}$.

Siano $U_1, U_2, \dots$ con determinazioni in $]0,1[$, iid, con $U_i \sim \text{Unif}(0,1)$.

1) $X_1, X_2, \dots$ stocasticamente indipendenti

Allora,

$$F_{X_1}^{\leftarrow}(U_1), F_{X_2}^{\leftarrow}(U_2), \dots$$

sono stocasticamente indipendenti e tali che $F_{X_i}^{\leftarrow}(U_i) \stackrel{d}{=} X_i$. Pertanto, $(F_{X_1}^{\leftarrow}(U_1), F_{X_2}^{\leftarrow}(U_2), \dots)$ simula $(X_1, X_2, \dots)$.

Se $U_1, U_2, \dots$ sono osservabili e $u_1, u_2, \dots$ i valori osservati, $x'_1 = F_{X_1}^{\leftarrow}(u_1), x'_2 = F_{X_2}^{\leftarrow}(u_2), \dots$ è una determinazione simulata di $X_1, X_2, \dots$.

2) $X_1, X_2, \dots$ non stocasticamente indipendenti

Si prova che, se $U_1, U_2, \dots$ sono osservabili e $u_1, u_2, \dots$ i valori osservati,

$$x'_1 = F_{X_1}^{\leftarrow}(u_1), x'_2 = F_{X_2|X_1=x'_1}^{\leftarrow}(u_2), x'_3 = F_{X_3|X_1=x'_1, X_2=x'_2}^{\leftarrow}(u_3), \dots$$

è una determinazione simulata di $X_1, X_2, \dots$.

Si noti che la simulazione avviene in modo sequenziale.

Nel caso del processo, da un punto di vista operativo, si ottiene un segmento iniziale di una traiettoria simulata.

In entrambi i casi, per ottenere m determinazioni simulate, si devono generare m sequenze indipendenti di numeri pseudo-casuali

$$\begin{array}{rcl} u_1^{(1)}, u_2^{(1)}, \dots & \rightarrow & \mathbf{x}'^{(1)} \\ & \dots & \\ u_1^{(m)}, u_2^{(m)}, \dots & \rightarrow & \mathbf{x}'^{(m)} \end{array}$$

ALCUNI ALGORITMI PER LA SIMULAZIONE

Presentiamo alcuni algoritmi per ottenere determinazioni simulate di alcune distribuzioni.

Nei software statistici vi sono algoritmi efficienti per molte distribuzioni di interesse applicativo.

Simulazione di un numero aleatorio con distribuzione normale

Sia $W \sim N(0,1)$.

Se U_1, U_2 sono iid con $U_i \sim \text{Unif}(0,1)$. Si prova che (Box-Müller transformation)

$$W_1 = \sqrt{-2\log U_1} \cos(2\pi U_2)$$

$$W_2 = \sqrt{-2\log U_1} \sin(2\pi U_2)$$

sono iid con distribuzione $N(0,1)$.

Simulazione: si genera una coppia di numeri pseudocasuali, u_1, u_2 , e si calcola $w_1 = \sqrt{-2\log u_1} \cos(2\pi u_2)$, $w_2 = \sqrt{-2\log u_1} \sin(2\pi u_2)$ ottenendo una coppia di determinazioni simulate dalla $N(0,1)$.

Sia $X \sim N(\mu, \sigma^2)$.

Si ha

$$W = \frac{X - \mu}{\sigma} \sim N(0,1) \quad \text{e} \quad X = \sigma W + \mu.$$

Allora, se W' simula $W \Rightarrow X' = \sigma W' + \mu$ simula X .

Simulazione: si genera una determinazione simulata w' dalla $N(0,1)$, ad esempio $w' = \sqrt{-2 \log u_1} \cos(2\pi u_2)$ con u_1, u_2 numeri pseudocasuali, e si calcola $x' = \sigma w' + \mu$.

Simulazione di un numero aleatorio con distribuzione log-normale

Sia $X \sim LN(\mu, \sigma)$.

Si ha

$$X \sim LN(\mu, \sigma) \Leftrightarrow Y = \log X \sim N(\mu, \sigma^2) \quad \text{e} \quad X = e^Y.$$

Allora, se Y' simula $Y \Rightarrow X' = e^{Y'}$ simula X .

Simulazione: si genera una determinazione simulata w' dalla $N(0,1)$, si calcola $y' = \sigma w' + \mu$ determinazione simulata di Y , e quindi si calcola $x' = e^{y'}$.

Simulazione di un numero aleatorio con distribuzione gamma

Sia $X \sim EXP(\rho)$.

Per $x > 0$, si ha $F_X(x) = 1 - e^{-\rho x}$, funzione strettamente crescente. L'inversa è $F_X^{-1}(u) = -\frac{1}{\rho} \log(1 - u)$, $u \in]0,1[$.

Allora, se $U \sim Unif(0,1) \Rightarrow X' = F_X^{-1}(U) = -\frac{1}{\rho} \log(1 - U)$ simula X .

Notiamo che $U \sim Unif(0,1) \Rightarrow 1 - U \sim Unif(0,1)$. Allora per simulare X si può considerare $X' = -\frac{1}{\rho} \log U$.

Simulazione: si genera un numero pseudocasuale u e si calcolare $x' = -\frac{1}{\rho} \log u$.

Sia $X \sim \text{gamma}(n, \rho)$, $n \in \mathbb{N}$, distribuzione erlangiana.

Si ha

$$X_1, \dots, X_n \text{ iid, con } X_i \sim \text{EXP}(\rho) \Rightarrow \sum_{i=1}^n X_i \sim \text{gamma}(n, \rho).$$

Allora, se $U_1, \dots, U_n$ iid con $U_i \sim \text{Unif}(0,1)$,

$$X'_1 = -\frac{1}{\rho} \log U_1, \dots, X'_n = -\frac{1}{\rho} \log U_n \text{ sono iid e } X'_i \sim \text{EXP}(\rho)$$

$$\Rightarrow X' = \sum_{i=1}^n X'_i = \sum_{i=1}^n -\frac{1}{\rho} \log U_i \text{ simula } X.$$

Simulazione: si generano $u_1, \dots, u_n$ numeri pseudocasuali e si calcola $x' = \sum_{i=1}^n -\frac{1}{\rho} \log u_i$.

Sia $X \sim \text{gamma}(\alpha, \rho)$.

Ci sono diversi algoritmi per la simulazione la cui efficacia dipende dal valore di α .

Simulazione di un numero aleatorio con distribuzione di Poisson

Sia $N \sim P(\lambda)$.

Può essere simulato in modo diretto come visto per i numeri aleatori discreti: $N' = F_N^{\leftarrow}(U)$, con F_N la funzione di ripartizione di N e $U \sim Unif(0,1)$.

Simulazione: si genera un numero pseudocasuale u , se $F_N(n-1) < u \leq F_N(n)$, $n' = F_N^{\leftarrow}(u) = n$.

Un algoritmo per simulare N si può ottenere ricorrendo ad alcuni risultati sul processo di Poisson.

Ricordiamo che se

$$\{N(t), t \geq 0\}$$

è un processo di Poisson di intensità λ , il processo dei tempi di interarrivo $\{W_1, W_2, \dots\}$ è un processo di numeri aleatori iid, con

$$W_n \sim EXP(\lambda).$$

Posto $T_n = W_1 + \dots + W_n$, il tempo di attesa per l'arrivo n -esimo, si ha

$$\begin{aligned}(N(t) = 0) &= (t < T_1) \\ &= (t < W_1),\end{aligned}$$

$$\begin{aligned}(N(t) = n) &= (T_n \leq t \wedge T_{n+1} > t) = (T_n \leq t < T_{n+1}) \\ &= (W_1 + \dots + W_n \leq t < W_1 + \dots + W_n + W_{n+1}), \text{ per } n \geq 1.\end{aligned}$$

Possiamo rappresentare $N(t) \sim P(\lambda t)$ con

$$N(t) = \sum_{n=0}^{+\infty} n |N(t) = n| = \sum_{n \geq 1} n |W_1 + \dots + W_n \leq t < W_1 + \dots + W_n + W_{n+1}|.$$

In particolare, $N(1) \sim P(\lambda)$ con

$$N(1) = \sum_{n=0}^{+\infty} n |N(1) = n| = \sum_{n \geq 1} n |W_1 + \dots + W_n \leq 1 < W_1 + \dots + W_n + W_{n+1}|.$$

Se $U_1, U_2, \dots$ iid con $U_i \sim \text{Unif}(0,1)$, allora

$$W'_1 = -\frac{1}{\lambda} \log U_1, W'_2 = -\frac{1}{\lambda} \log U_2, \dots$$

simula $\{W_1, W_2, \dots\}$.

Segue che

$$N' = \sum_{n \geq 1} n |W'_1 + \dots + W'_n \leq 1 < W'_1 + \dots + W'_n + W'_{n+1}|$$

simula $N(1)$ e N :

$$N' = 0 \Leftrightarrow W'_1 > 1 \Leftrightarrow -\frac{1}{\lambda} \log U_1 > 1 \Leftrightarrow \log U_1 < -\lambda$$

$$\Leftrightarrow U_1 < e^{-\lambda},$$

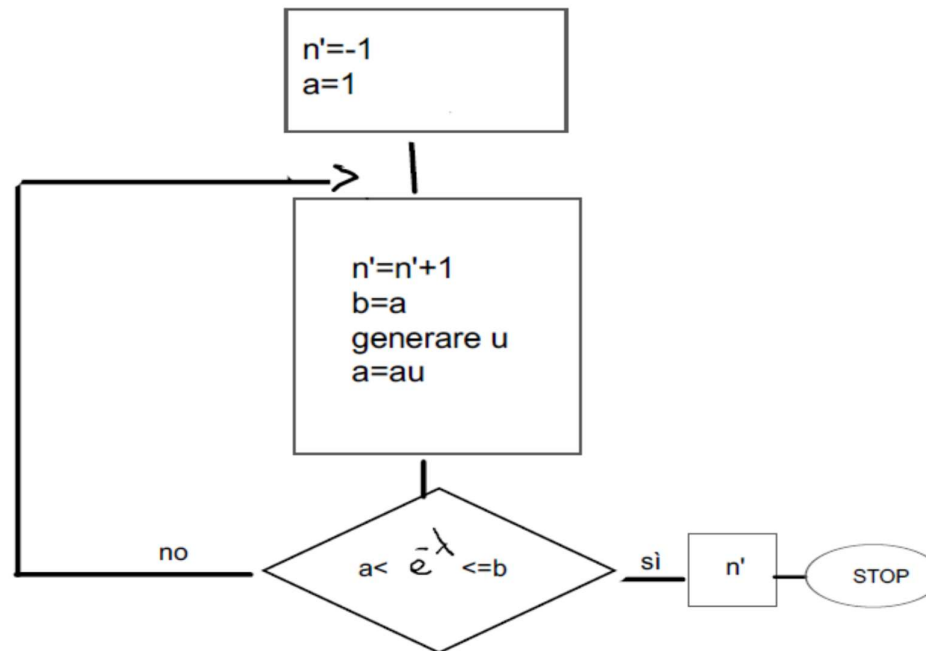
$$N' = n \Leftrightarrow W'_1 + \dots + W'_n \leq 1 < W'_1 + \dots + W'_n + W'_{n+1} \Leftrightarrow \sum_{i=1}^n -\frac{1}{\lambda} \log U_i \leq 1 < \sum_{i=1}^{n+1} -\frac{1}{\lambda} \log U_i$$

$$\Leftrightarrow \sum_{i=1}^n \log U_i \geq -\lambda > \sum_{i=1}^{n+1} \log U_i$$

$$\Leftrightarrow \prod_{i=1}^{n+1} U_i < e^{-\lambda} \leq \prod_{i=1}^n U_i.$$

Simulazione: si può procedere in modo sequenziale

- (1) si genera un numero pseudocasuale u_1 , se $\nearrow u_1 < e^{-\lambda}$ $\Rightarrow n' = 0$
 $\searrow$ altrimenti (2)
- (2) si genera un numero pseudocasuale u_2 , se $\nearrow u_1 u_2 < e^{-\lambda} (\leq u_1)$ $\Rightarrow n' = 1$
 $\searrow$ altrimenti (3)
- (3) si genera un numero pseudocasuale u_3 , se $\nearrow u_1 u_2 u_3 < e^{-\lambda} (\leq u_1 u_2)$ $\Rightarrow n' = 2$
 $\searrow$ altrimenti (4)
- (4) ...



Simulazione di un numero aleatorio con distribuzione binomiale negativa

Sia $N \sim BN(\alpha, p)$.

La distribuzione binomiale negativa si può ottenere come mistura di poissoniane con misturante gamma. Siano N' e Θ con

- $N' | \Theta = \theta \sim P(\lambda\theta)$,
- $\Theta \sim \text{gamma}(\alpha, \alpha)$.

Si ha,
$$N' \sim BN\left(\alpha, \frac{\alpha}{\alpha + \lambda}\right), \quad \text{con} \quad E(N') = \lambda, \text{ var}(N') = \lambda + \alpha^{-1}\lambda^2.$$

Dalla $\frac{\alpha}{\alpha + \lambda} = p$, si ricava λ . Per simulare N si può considerare la coppia (Θ, N') e applicare il procedimento visto per la simulazione di vettori aleatori con componenti non indipendenti.

Simulazione:

- si genera un numero pseudocasuale u_1 e si calcola $\theta' = F_{\Theta}^{\leftarrow}(u_1)$ ovvero si genera una determinazione simulata dalla $\text{gamma}(\alpha, \alpha)$,
- si genera un numero pseudocasuale u_2 e si calcola $n' = F_{N' | \Theta = \theta'}^{\leftarrow}(u_2)$ ovvero si genera una determinazione simulata dalla $P(\lambda\theta')$, n' è determinazione simulata di N .

Simulazione di un numero aleatorio con distribuzione Poisson composta

Sia $X = \sum_{i=1}^N Y_i \sim \text{Poisson composta}(\lambda, F_Y)$.

Metodo diretto

Si basa sull'osservazione che X è funzione di un processo stocastico con componenti non stocasticamente indipendenti

$$X = \varphi(N, Y_1, Y_2, \dots).$$

Simulazione:

- (1) si genera un numero pseudocasuale u_0 e si calcola $n' = F_N^{\leftarrow}(u_0)$ ovvero si genera una determinazione simulata dalla $P(\lambda)$; se $n' = 0$ allora $x' = 0$ è la determinazione simulata di X , altrimenti si va al (2),
- (2) poiché $Y_1|N = n', \dots, Y_{n'}|N = n'$ sono iid con funzione di ripartizione F_Y , si generano $u_1, \dots, u_{n'}$ e si calcolano $y'_1 = F_Y^{\leftarrow}(u_1), \dots, y'_{n'} = F_Y^{\leftarrow}(u_{n'})$ ovvero si generano n' determinazioni simulate dalla F_Y ,
- (3) si calcola $x' = y'_1 + \dots + y'_{n'}$, x' è determinazione simulata di X .

Generando più determinazioni simulate, si può calcolare la funzione di ripartizione empirica che è una stima della F_X , la funzione di ripartizione di X .

Metodo basato sull'approssimazione Wilson-Hilferty

Siano F_X la funzione di ripartizione di X , $\mu_X = E(X)$, $\sigma_X = \sqrt{\text{var}(X)}$, $\gamma_X = E[(X - \mu_X)^3]/\sigma_X^3$. Si ha che

$$F_X(x) \cong \phi\left(h\left(\frac{x-\mu_X}{\sigma_X}\right)\right), \quad x \geq 0, \quad (*)$$

dove

$$h(t) = c_1 + c_2(t + c_3)^{1/3}$$

con

$$c_1 = \frac{1}{6}\gamma_X - \frac{6}{\gamma_X}, \quad c_2 = 3\left(\frac{2}{\gamma_X}\right)^{2/3}, \quad c_3 = \frac{2}{\gamma_X},$$

e ϕ è la funzione di ripartizione della $N(0,1)$.

La (*) è detta approssimazione Wilson-Hilferty. E' stata determinata per approssimare la funzione di ripartizione della gamma. Può essere usata per approssimare la funzione di ripartizione della Poisson composta se γ_X è "piccolo" ($\gamma_X \leq 1$).

Per (*), se $W \sim N(0,1)$, essendo h crescente, si ha

$$\begin{aligned} F_X(x) &\cong \phi\left(h\left(\frac{x - \mu_X}{\sigma_X}\right)\right) = Pr\left(W \leq h\left(\frac{x - \mu_X}{\sigma_X}\right)\right) = Pr\left(h^{-1}(W) \leq \frac{x - \mu_X}{\sigma_X}\right) \\ &= Pr(\sigma_X h^{-1}(W) + \mu_X \leq x). \end{aligned}$$

Allora, $X' = \sigma_X h^{-1}(W) + \mu_X$

ha approssimativamente la stessa funzione di ripartizione X e può essere utilizzato per simulare X .

Il numero aleatorio X' dipende solo dai primi tre momenti della distribuzione Poisson composta che si calcolano facilmente:

$$\mu_X = E(N)E(Y), \quad \sigma_X = \sqrt{E(N)E(Y^2)}, \quad \gamma_X = E(Y^3)/\sqrt{E(N)E(Y^2)^3}.$$

Simulazione:

- si genera una determinazione simulata w' dalla $N(0,1)$,
- si calcola $x' = \sigma_X h^{-1}(w') + \mu_X$.

Metodo basato sull'approssimazione Normal Power

Siano F_X la funzione di ripartizione di X , $\mu_X = E(X)$, $\sigma_X = \sqrt{\text{var}(X)}$, $\gamma_X = E[(X - \mu_X)^3]/\sigma_X^3$. Si ha che

$$F_X(x) \cong \phi\left(-\frac{3}{\gamma_X} + \sqrt{\frac{9}{\gamma_X^2} + 1 + \frac{6}{\gamma_X} \frac{x - \mu_X}{\sigma_X}}\right), \quad x \geq 0, \quad (**)$$

con ϕ la funzione di ripartizione della $N(0,1)$.

La (**) è detta approssimazione Normal Power. Può essere usata per approssimare la funzione di ripartizione della Poisson composta se γ_X è non troppo elevato.

Si ricava che se $W \sim N(0,1)$,

$$X' = \frac{\sigma_X \gamma_X}{6} (W^2 - 1) + \sigma_X W + \mu_X$$

ha approssimativamente la stessa funzione di ripartizione X e può essere utilizzato per simulare X .

Simulazione:

- si genera una determinazione simulata w' dalla $N(0,1)$,
- si calcola $x' = \frac{\sigma_X \gamma_X}{6} (w'^2 - 1) + \sigma_X w' + \mu_X$.

Osservazione. Se X rappresenta il risarcimento totale per un portafoglio, in un fissato periodo di tempo, il capitale di rischio basato sul VaR al livello p è

$$VaR[X; p] - P.$$

Si deve calcolare $VaR[X; p] = F_X^{\leftarrow}(p)$. Se a X si assegna distribuzione Poisson composta, la F_X non è ottenibile in forma chiusa.

Si può approssimare il VaR tramite approccio simulativo:

- si generano m determinazioni simulate di X , $x'_1, \dots, x'_m$,
- si approssima la F_X con la funzione di ripartizione empirica ottenuta dalle determinazioni simulate, $F_m(x) = \frac{\text{card}\{i \mid x'_i \leq x\}}{m}$, $x \in \mathbb{R}$,
- si approssima $VaR[X; p] \cong F_m^{\leftarrow}(p)$.

I quantili della funzione della ripartizione empirica sono i valori della statistica d'ordine. Posto

$$x'_{(1)} < \cdots < x'_{(m)},$$

se p è tale che

$$F_m(x'_{(h-1)}) < p \leq F_m(x'_{(h)}) \Leftrightarrow \frac{h-1}{m} < p \leq \frac{h}{m} \Leftrightarrow h-1 < mp \leq h$$
$$\Rightarrow F_m^{\leftarrow}(p) = x'_{(h)}$$

$F_m^{\leftarrow}(p)$ è il valore della statistica d'ordine $x'_{(h)}$ con h il minimo intero $\geq mp$.

Ad esempio, se $p = 0.995$ e $m = 10000$, allora $F_m^{\leftarrow}(p) = x'_{(9950)} \cong VaR[X; p]$.

Simulazione di un processo di Poisson

Benché sia a parametro continuo, per il processo di Poisson si possono ottenere traiettorie iniziali simulate. Ciò segue dal fatto che il processo può essere ottenuto a partire dal processo dei tempi di interarrivo $\{W_1, W_2, \dots\}$ che è un processo di numeri aleatori iid, con

$$W_n \sim \text{EXP}(\lambda).$$

Fissato $\tau > 0$, vediamo come ottenere una traiettoria simulata di $\{N(t), 0 \leq t \leq \tau\}$.

Ricordiamo che

$$(N(t) = 0) = (t < T_1),$$

$$(N(t) = n) = (T_n \leq t < T_{n+1}), \text{ per } n \geq 1,$$

dove $T_n = W_1 + \dots + W_n$, e che se $U_1, U_2, \dots$ iid con $U_i \sim \text{Unif}(0,1)$, allora il processo

$$W'_1 = -\frac{1}{\lambda} \log U_1, W'_2 = -\frac{1}{\lambda} \log U_2, \dots$$

simula $\{W_1, W_2, \dots\}$.

Simulazione: si può procedere in modo sequenziale

(1) si genera una determinazione simulata w'_1 di W_1 e si calcola la determinazione simulata

di T_1 , $t'_1 = w'_1$; se $\nearrow t'_1 > \tau$ $n'(t) = 0, t \in [0, \tau] \Rightarrow \text{STOP}$
 $\searrow \text{altrimenti}$ $n'(t) = 0, t \in [0, t'_1[\Rightarrow (2)$

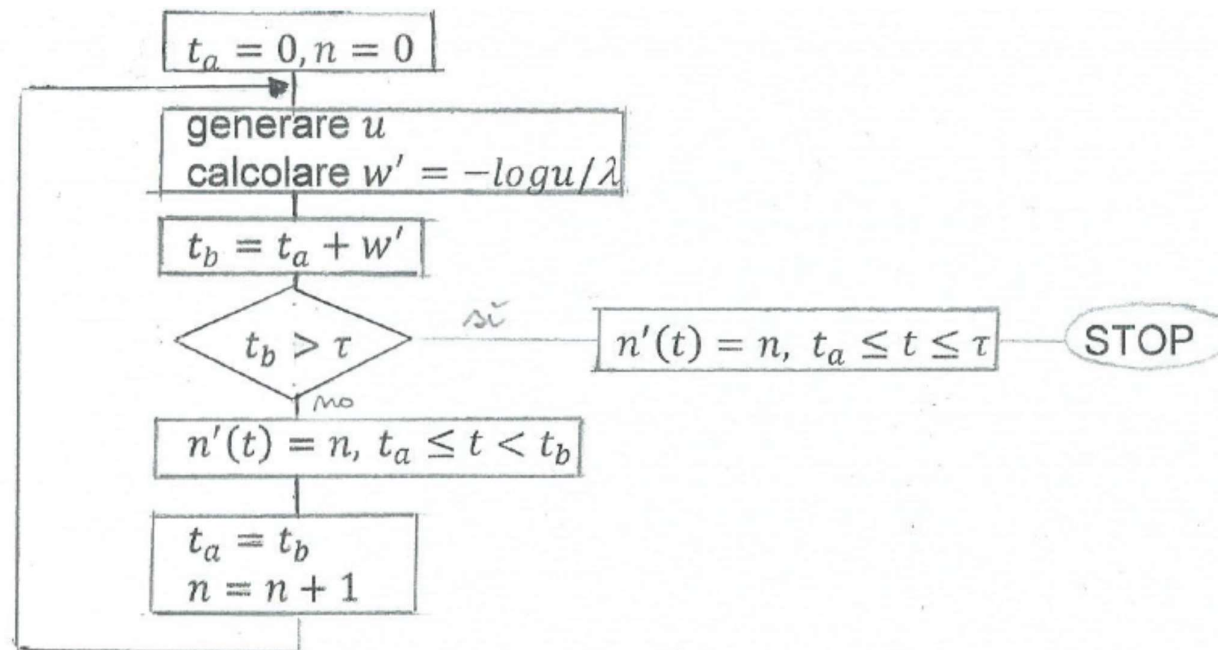
(2) si genera una determinazione simulata w'_2 di W_2 e si calcola la determinazione simulata

di T_2 , $t'_2 = t'_1 + w'_2$; se $\nearrow t'_2 > \tau$ $n'(t) = 1, t \in [t'_1, \tau] \Rightarrow \text{STOP}$
 $\searrow \text{altrimenti}$ $n'(t) = 1, t \in [t'_1, t'_2[\Rightarrow (3)$

(3) si genera una determinazione simulata w'_3 di W_3 e si calcola la determinazione simulata

di T_3 , $t'_3 = t'_2 + w'_3$; se $\nearrow t'_3 > \tau$ $n'(t) = 2, t \in [t'_2, \tau] \Rightarrow \text{STOP}$
 $\searrow \text{altrimenti}$ $n'(t) = 2, t \in [t'_2, t'_3[\Rightarrow (4)$

(4) ...



Valutazione della probabilità di rovina mediante simulazione

Ricordiamo che nel modello classico della Teoria collettiva del rischio, il processo di rischio è $\{R(t), t \geq 0\}$ con

$$R(t) = R + ct - S(t),$$

dove R è capitale allocato, ct sono i premi incassati in $[0, t]$, $S(t)$ è il risarcimento cumulato in $[0, t]$,

$$S(t) = \sum_{h=1}^{N(t)} Y_h$$

dove

- $N(t)$ è il numero di sinistri che colpiscono il portafoglio in $[0, t]$,
- Y_h è l'importo del risarcimento per il sinistro h -esimo,

con la convenzione $S(t) = 0$, se $N(t) = 0$.

Nelle ipotesi del modello, $\{S(t), t \geq 0\}$ è un processo Poisson composto (λ, F_Y) .

La **probabilità di rovina nell'intervallo limitato** $[0, \tau]$, **in tempo continuo**, a partire dal capitale iniziale R ,

$$\psi(R, \tau) = Pr(V_{0 \leq t \leq \tau}(R(t) < 0)).$$

Tale probabilità può essere facilmente stimata mediante simulazione.

Si possono infatti simulare più traiettorie del processo $\{R(t), 0 \leq t \leq \tau\}$ e stimare $\psi(R, \tau)$ tramite la frequenza relativa delle traiettorie che portano alla rovina

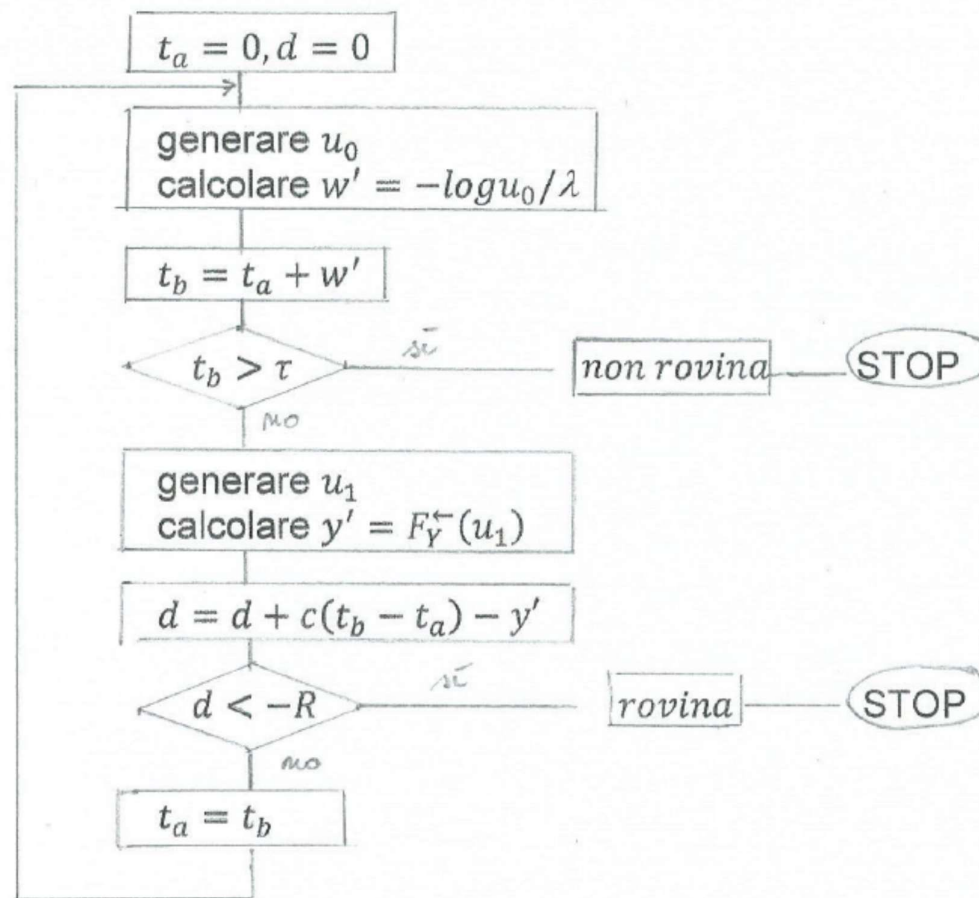
$$\hat{\psi}(R, \tau) = \frac{\text{nr. traiettorie che portano alla rovina}}{\text{nr. traiettorie simulate}}.$$

Vediamo come simulare una traiettoria.

Osserviamo che la rovina può verificarsi solo nell'istante in cui si verifichi un sinistro. Allora, per le ipotesi del modello, si possono simulare gli istanti nei quali si verificano i sinistri, come visto nella simulazione di un processo di Poisson, e le entità dei corrispondenti risarcimenti.

Poniamo $D(t) = ct - S(t)$. Si ha $R(t) = R + D(t)$ e

$$R(t) < 0 \Leftrightarrow D(t) < -R.$$



Alcuni riferimenti

C.D. Daykin, T. Pentikäinen, M. Pesonen (1994), *Practical risk theory for actuaries*, Chapman & Hall

M. Denuit, A. Charpentier (2004), *Mathématiques de l'assurance non-vie*, Economica