

# Tecnica attuariale danni: Esercizi

Giada Cesaro

22/7/2020

## Indice

<b>1</b>	<b>Simulazioni</b>	<b>2</b>
1.1	Distribuzione uniforme . . . . .	2
1.2	Distribuzione normale . . . . .	3
1.2.1	L'algoritmo di Box-Muller . . . . .	3
1.2.2	La funzione rnorm . . . . .	4
1.2.3	Confronto degli algoritmi per simulazione di una normale . . . . .	6
1.3	Distribuzione lognormale . . . . .	6
1.4	Distribuzione esponenziale . . . . .	8
1.4.1	Simulazione usando la f.d.r inversa . . . . .	8
1.4.2	Simulazione usando rexp() . . . . .	8
1.4.3	Confronto dei due metodi . . . . .	9
1.4.4	Simulazioni ripetute per n=100 . . . . .	10
1.5	Distribuzione Gamma . . . . .	10
1.6	Distribuzione di Poisson . . . . .	12
1.6.1	Simulazione mediante l'algoritmo . . . . .	12
1.6.2	Usando la funzione inversa (a gradini) . . . . .	13
1.6.3	Confronto dei due metodi . . . . .	13
1.7	Distribuzione Poisson-composta . . . . .	14
1.7.1	Metodo diretto . . . . .	14
1.7.2	Approssimazione Wilson-Hilferty . . . . .	15
1.7.3	Approssimazione Normal Power . . . . .	15
1.7.4	Confronto metodo diretto, Wilson Hilferty e Normal Power . . . . .	16
<b>2</b>	<b>Probabilità di rovina</b>	<b>18</b>
2.1	Probabilità di rovina al variare di Theta . . . . .	18
2.2	Probabilità di rovina al variare di $F_y$ . . . . .	20
2.3	Probabilità di rovina al variare di Theta e R . . . . .	22
2.4	Probabilità di rovina al variare di Lambda . . . . .	23
<b>3</b>	<b>Metodi deterministici per il calcolo delle riserve sinistri</b>	<b>25</b>
3.1	Chain ladder . . . . .	25
3.2	Metodo di separazione (o di Taylor) . . . . .	27
3.3	Bornhuetter-Ferguson . . . . .	29
3.4	Fisher-Lange: metodo 2 . . . . .	30
3.5	Confronto dei metodi . . . . .	32

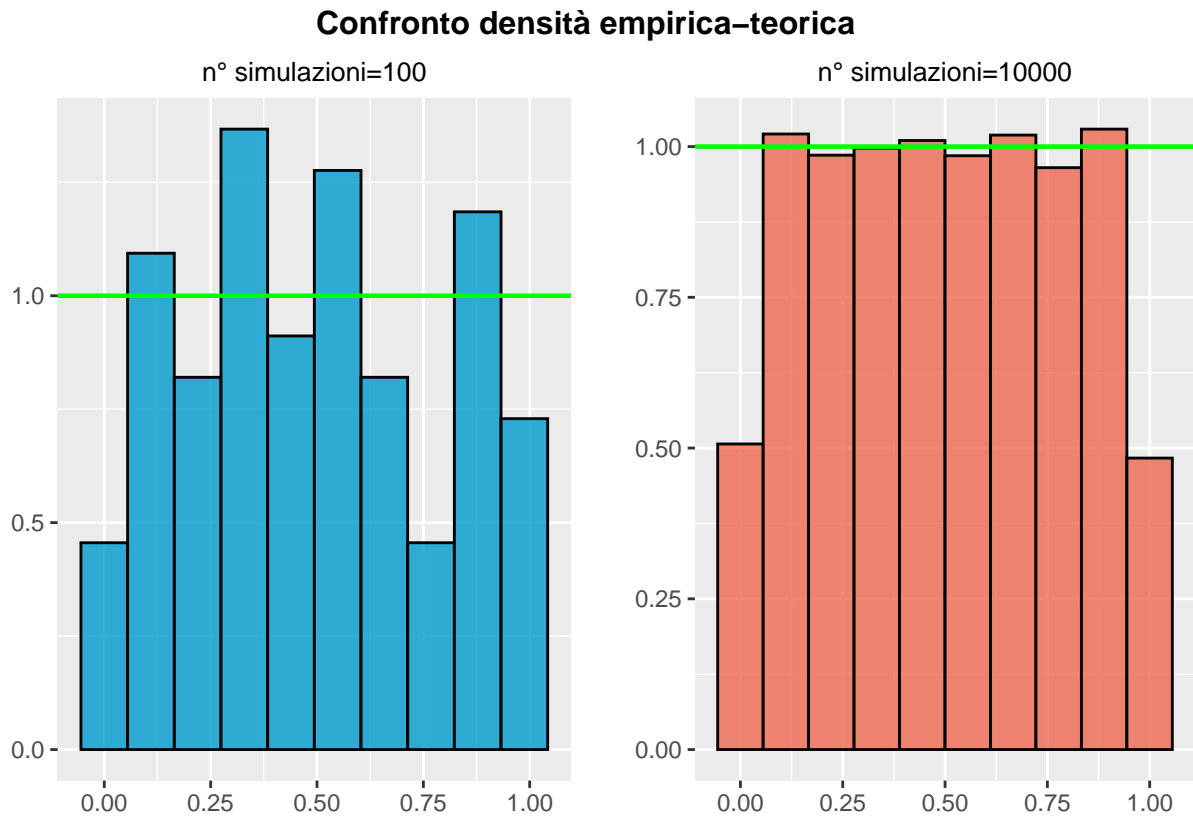
# 1 Simulazioni

**Consegna:** Per ciascuna delle seguenti distribuzioni eseguire un numero più basso ed uno più elevato di simulazioni e confrontare i seguenti elementi empirici e teorici: valori attesi, varianze, funzioni di ripartizioni e di densità. Nelle simulazioni usare gli algoritmi descritti a lezione e le funzioni disponibili in R.

Nelle seguenti simulazioni si è scelto di confrontare i risultati ottenuti con  $n = 100$  e  $n = 10000$  simulazioni. In nero sono rappresentate le funzioni di densità/probabilità e ripartizione reali per le varie distribuzioni, mentre quelle empiriche sono rappresentate in azzurro e corallo per  $n = 100$  e  $n = 10000$  rispettivamente.

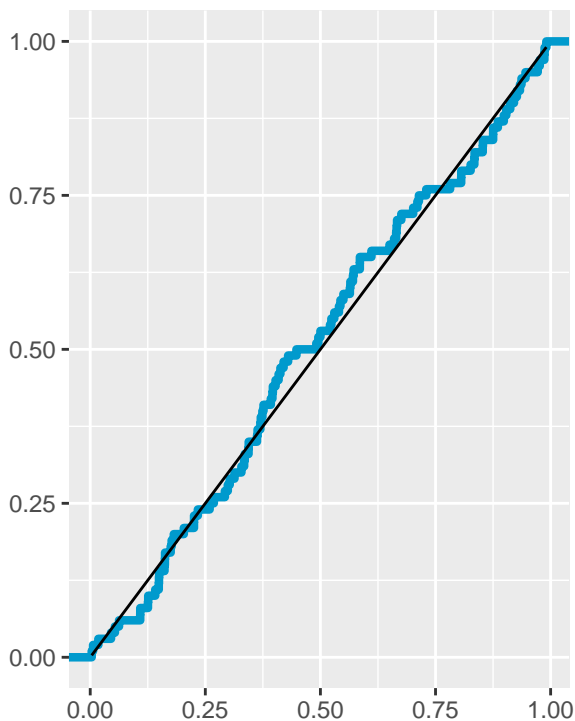
## 1.1 Distribuzione uniforme

Si simula da una  $Unif(0, 1)$  usando la funzione `runif()` di R.

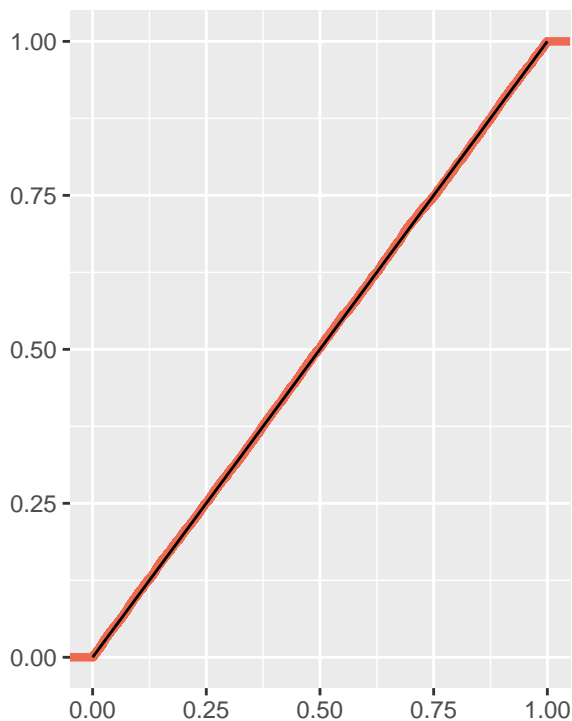


## Confronto funzione di ripartizione empirica–teorica

n° simulazioni= 100



n° simulazioni= 10000



La media teorica è  $\frac{a+b}{2}$ . Trattandosi di un'uniforme in  $[0,1]$  ( $a = 0$ ,  $b = 1$ ) si ha che la media è  $\frac{1}{2}$ , mentre la varianza è  $\frac{(b-a)^2}{12}$ , ovvero  $\frac{1}{12}$ .

Si rappresentano nella seguente tabella le medie e varianze empiriche:

	Media	Varianza
n=100	0.496	0.083
n=10000	0.499	0.084

## 1.2 Distribuzione normale

### 1.2.1 L'algoritmo di Box-Muller

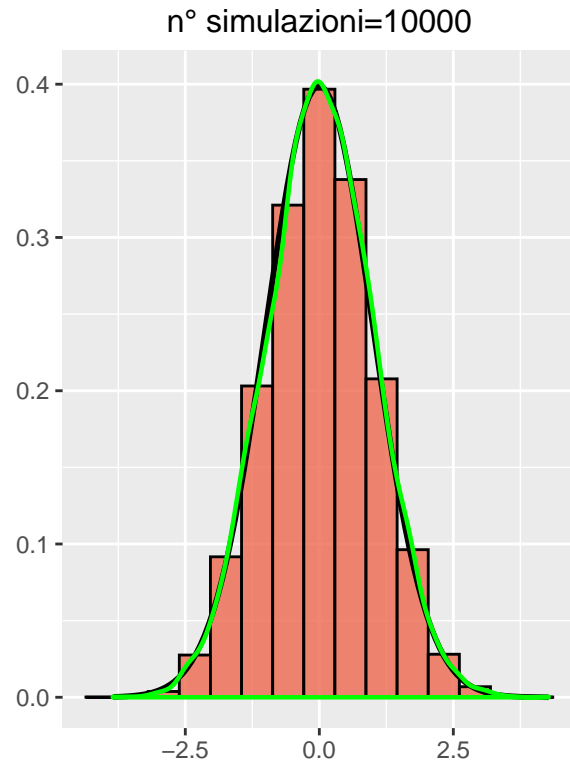
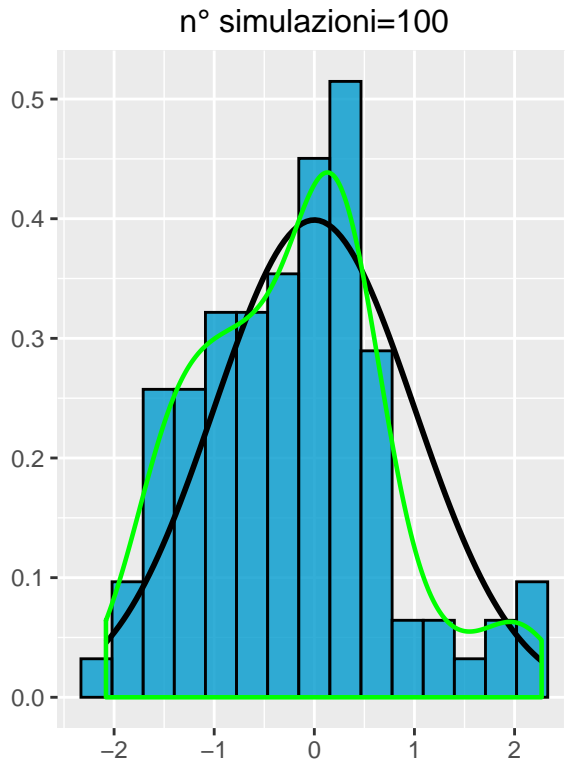
L'algoritmo permette di simulare da una  $N(0,1)$  generando valori da una  $Unif(0,1)$  e trasformandoli nel seguente modo:

$$U_1, U_2 \sim Unif(0,1) \text{ e ind.}$$

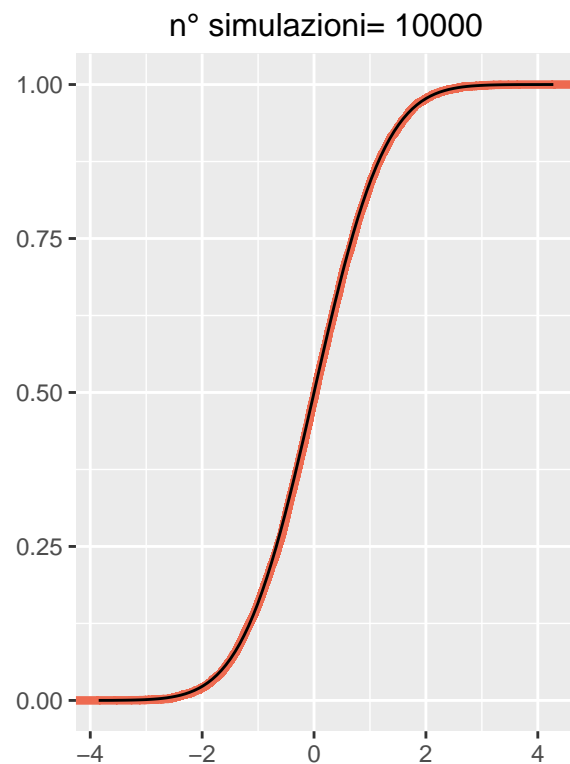
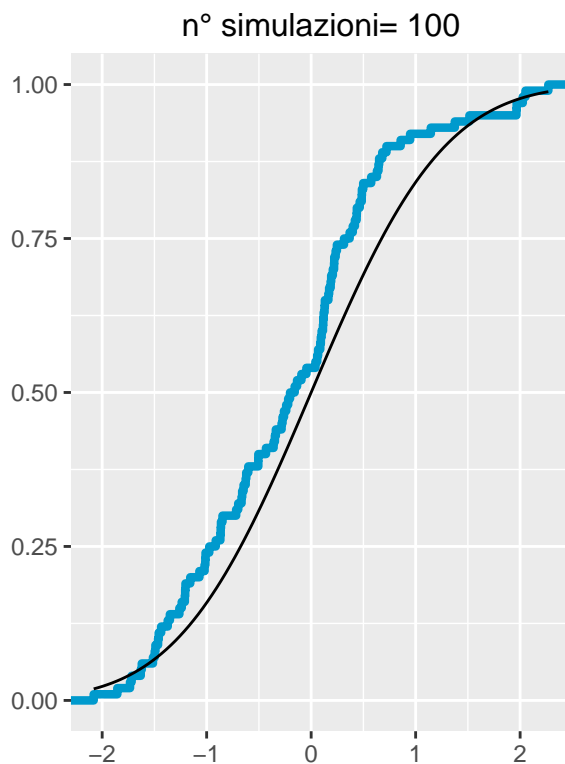
$$W_1 = \sqrt{-2\log(U_1)} \cos(2\pi U_2); \quad W_2 = \sqrt{-2\log(U_1)} \sin(2\pi U_2) \rightarrow W_1, W_2 \sim N(0,1) \text{ e ind.}$$

Nei grafici di confronto delle densità empiriche e teoriche, in verde si rappresenta la *smoothed density*.

### Confronto densità empirica–teorica



### Confronto funzione di ripartizione empirica–teorica

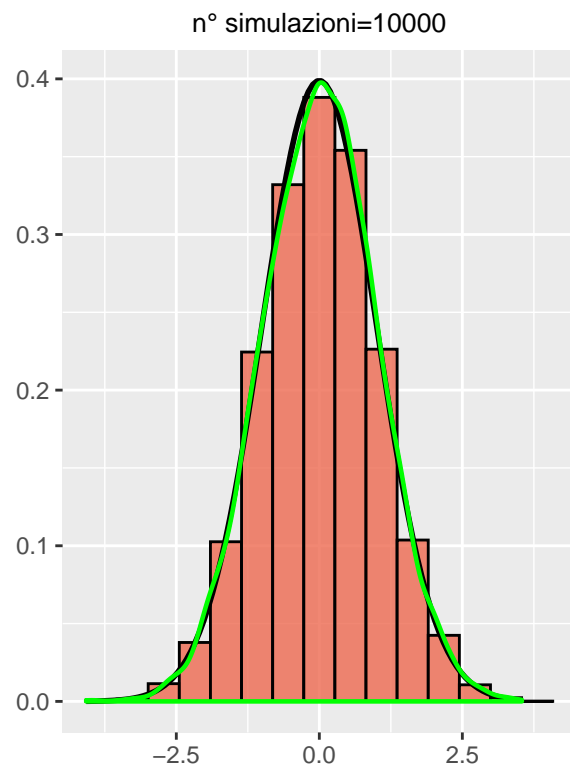
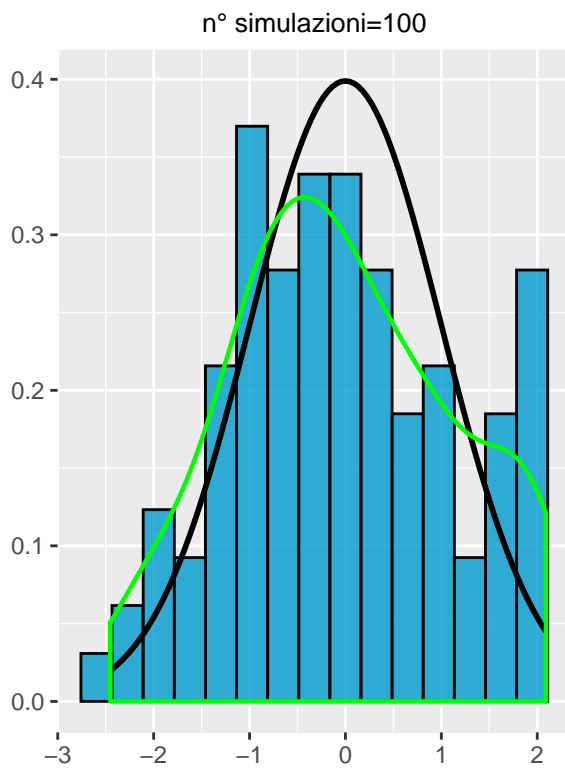


#### 1.2.2 La funzione `rnorm`

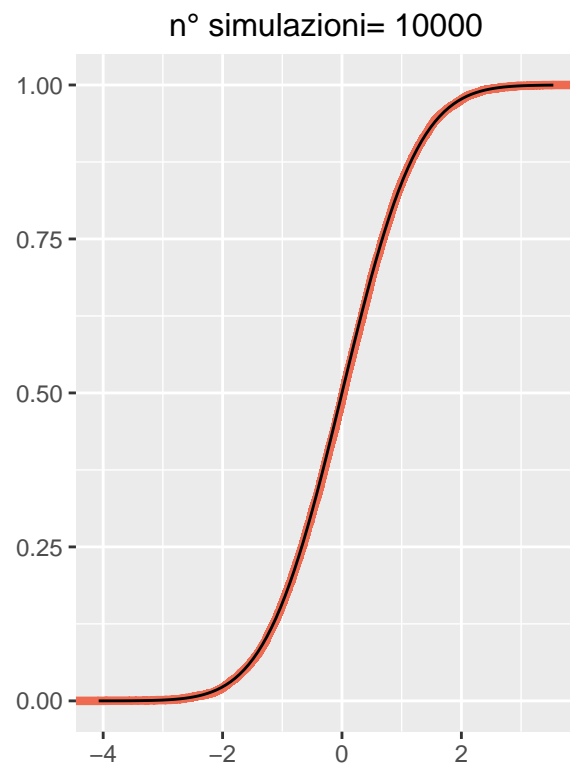
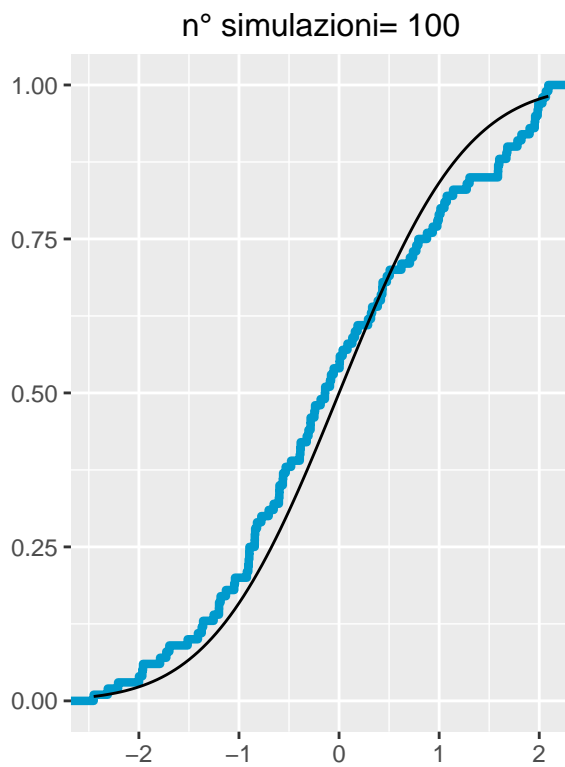
Si simula ora da una  $N(0, 1)$  usando la funzione `rnorm()` di R, che di default utilizza l'algoritmo di *Mersenne-Twister*. È possibile cambiare l'algoritmo di default tramite il parametro `RNG`. Si confrontano, poi, medie e varianze dei due

diversi algoritmi.

### Confronto densità empirica–teorica



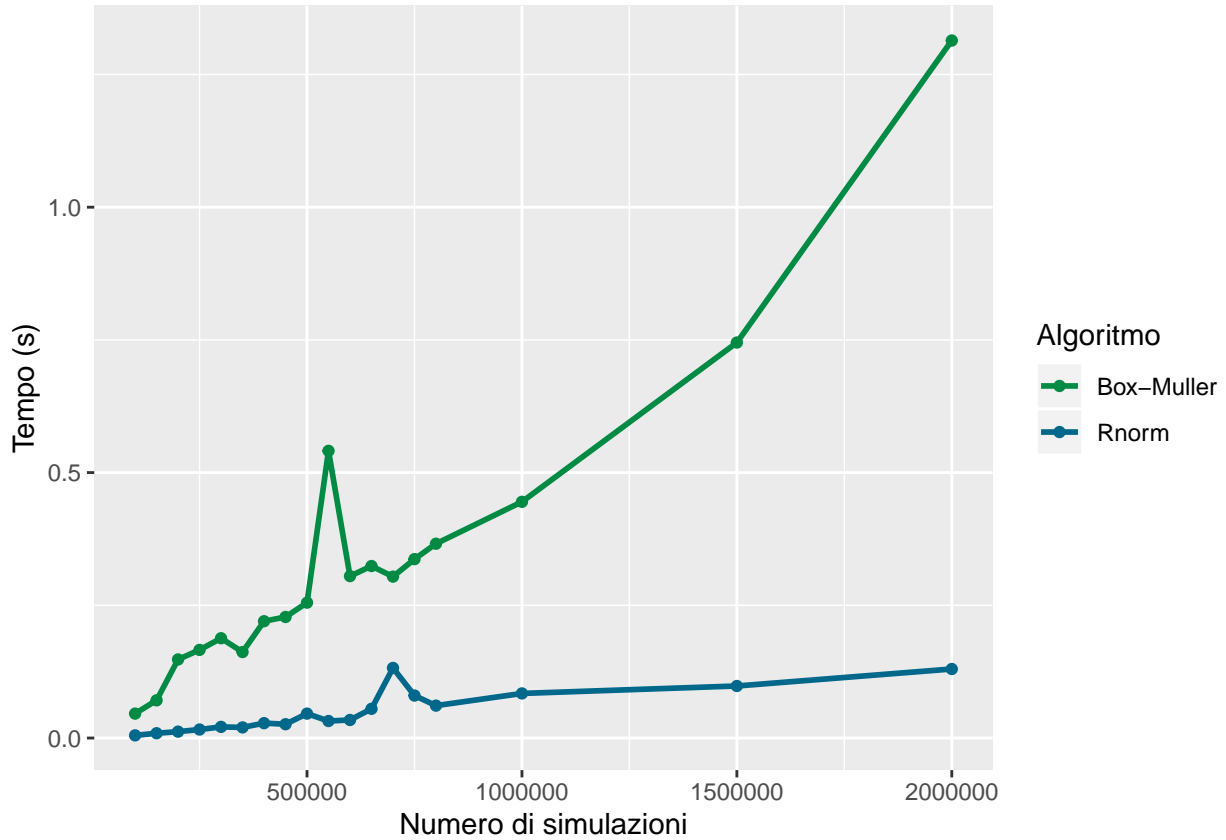
### Confronto funzione di ripartizione empirica–teorica



	m-BoxM	var-BoxM	m-rnorm	var-rnorm
n=100	-0.214	0.881	-0.031	1.346
n=10000	0.020	0.993	0.012	0.996

### 1.2.3 Confronto degli algoritmi per simulazione di una normale

Si confronta il tempo impiegato, per diversi numeri di simulazioni, utilizzando l'algoritmo di *Box-Muller* e la funzione `rnorm()`.



## 1.3 Distribuzione lognormale

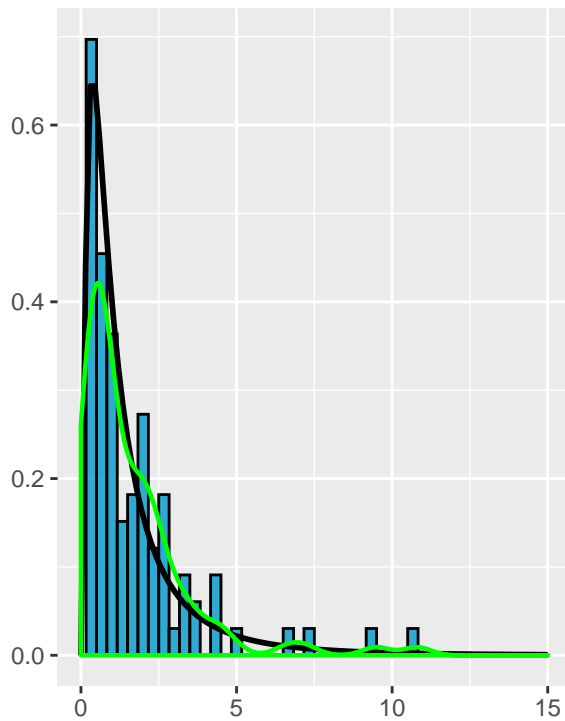
Se  $X \sim LN(\mu, \sigma)$  allora  $Y = \log(X) \sim N(\mu, \sigma^2)$ . L'idea è quindi di simulare  $Y'$  da una  $N(0, 1)$ , calcolare  $Y = \sigma Y' + \mu \sim N(\mu, \sigma^2)$  e ricavare la determinazione simulata di  $X$ ,  $x$ , tramite la determinazione simulata  $y$  di  $Y$  nel seguente modo:

$$x = e^y (= e^{\log(x)})$$

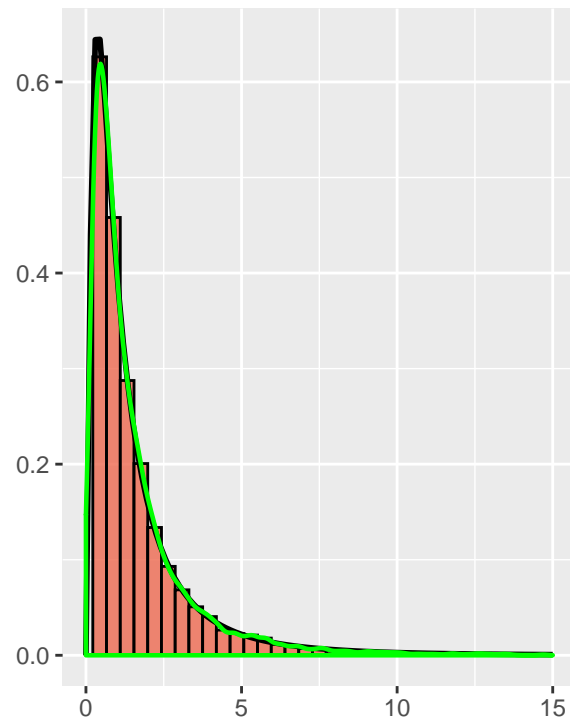
Nel seguente esempio si è scelto di simulare da una  $LN(0, 1)$ .

## Confronto densità empirica-teorica

n° simulazioni=100

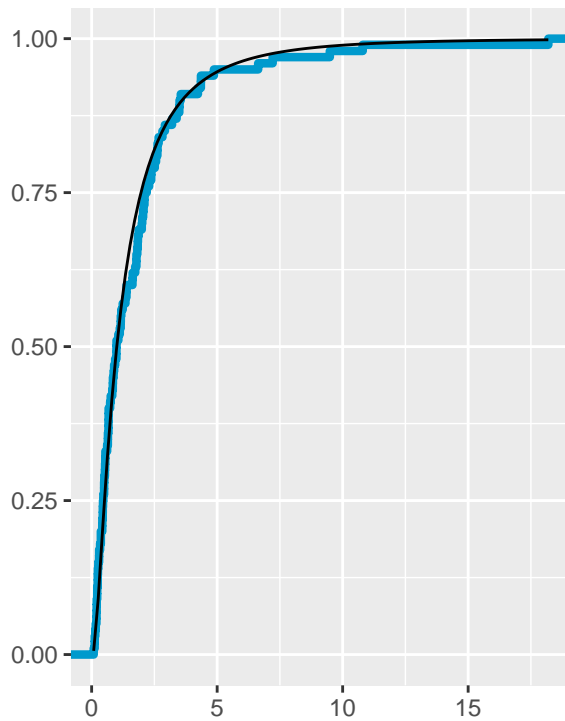


n° simulazioni=10000

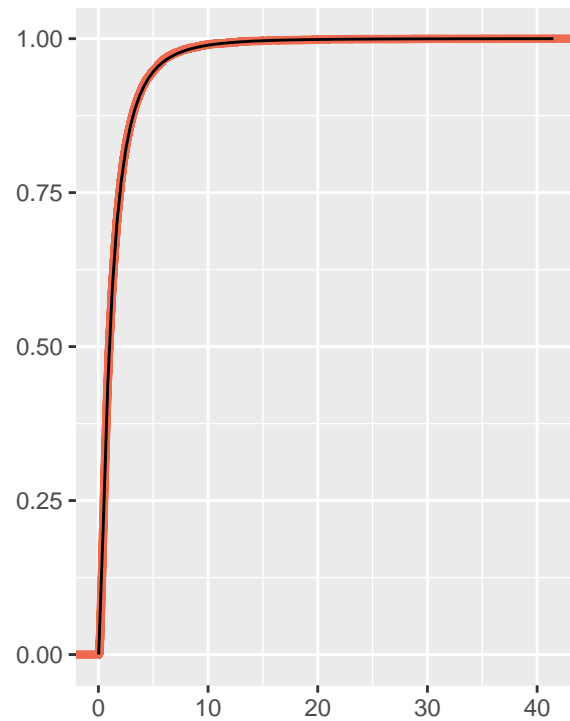


## Confronto funzione di ripartizione empirica-teorica

n° simulazioni= 100



n° simulazioni= 10000



La media e la varianza teoriche sono  $E(X) = e^{\mu + \frac{\sigma^2}{2}} = 1.6487213$  e  $Var(X) = e^{2\mu + \sigma^2}(e^{\sigma^2} - 1) = 4.6707743$ . Nella seguente tabella si mostrano  $\hat{\mu}$  e  $\hat{\sigma}^2$  per  $n = 100$  e  $n = 10000$ .

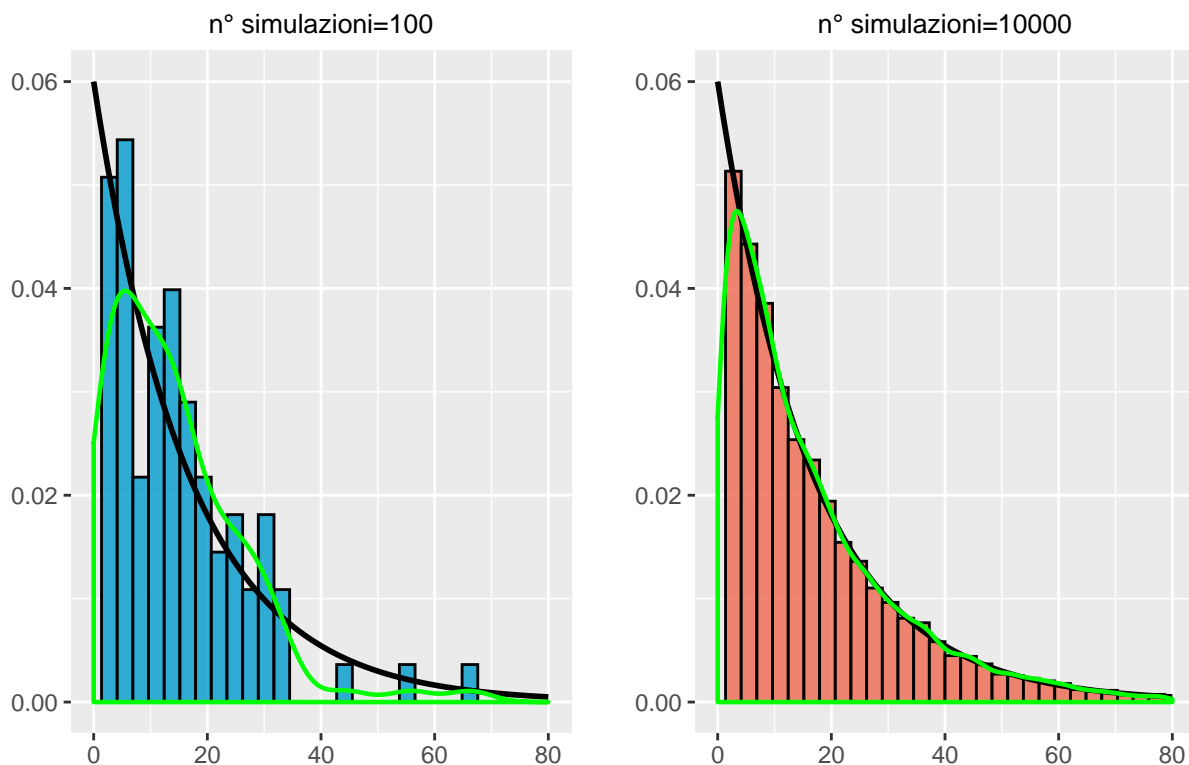
	Media	Varianza
n=100	1.797	6.101
n=10000	1.627	4.428

## 1.4 Distribuzione esponenziale

### 1.4.1 Simulazione usando la f.d.r inversa

Essendo la funzione di una variabile  $X \sim \text{Exp}(\rho)$  pari a  $F(x) = 1 - e^{-\rho x}$ , si può ricavare la funzione inversa. Risulta che  $X' = -\frac{1}{\rho} \log(U)$ , dove  $U \sim \text{Unif}(0, 1)$ . Le determinazioni simulate  $u_1, \dots, u_n$  permettono quindi di ricavare  $n$  determinazioni simulate di un'esponenziale con parametro  $\rho$  calcolando  $x'_1 = -\frac{1}{\rho} \log(u_1)$ , ...,  $x'_n = -\frac{1}{\rho} \log(u_n)$ .

### Confronto densità empirica-teorica

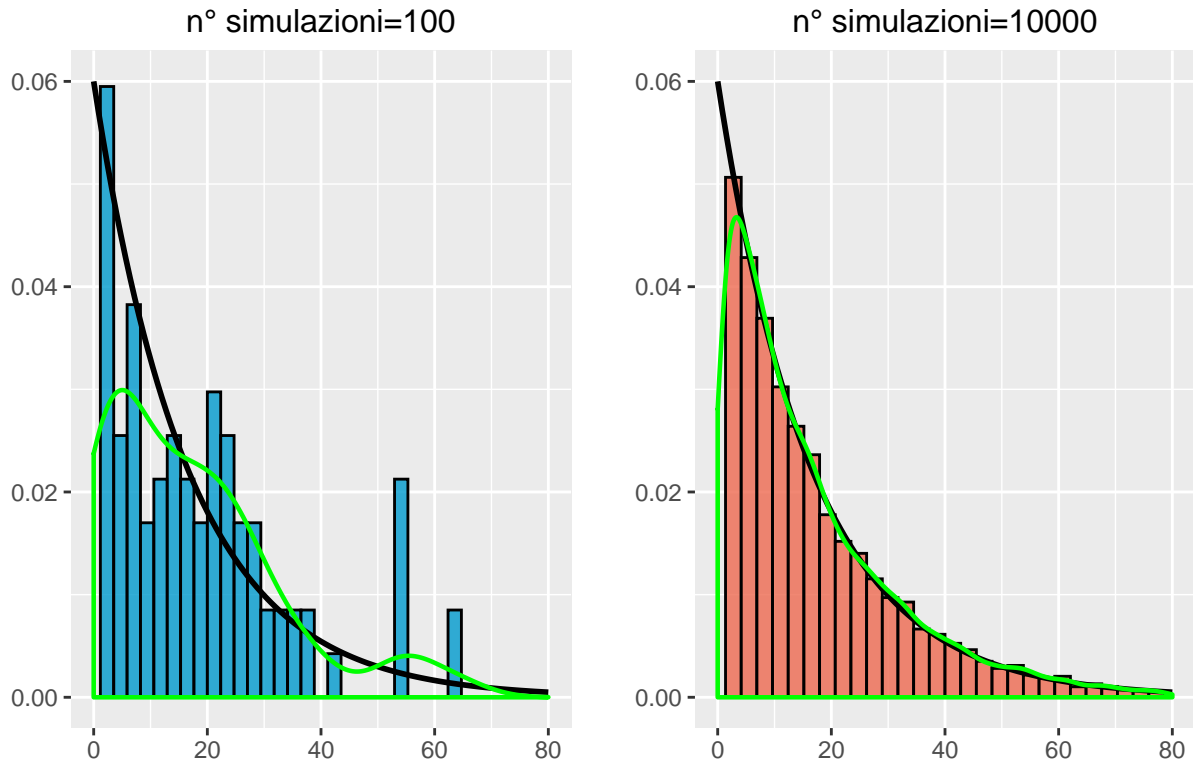


### 1.4.2 Simulazione usando rexp()

Si utilizza la funzione di `rexp()` già implementata in R per simulare da un'esponenziale con parametro  $\rho = 0.06$ , e con media  $\frac{1}{0.06} \approx 16.66$ .



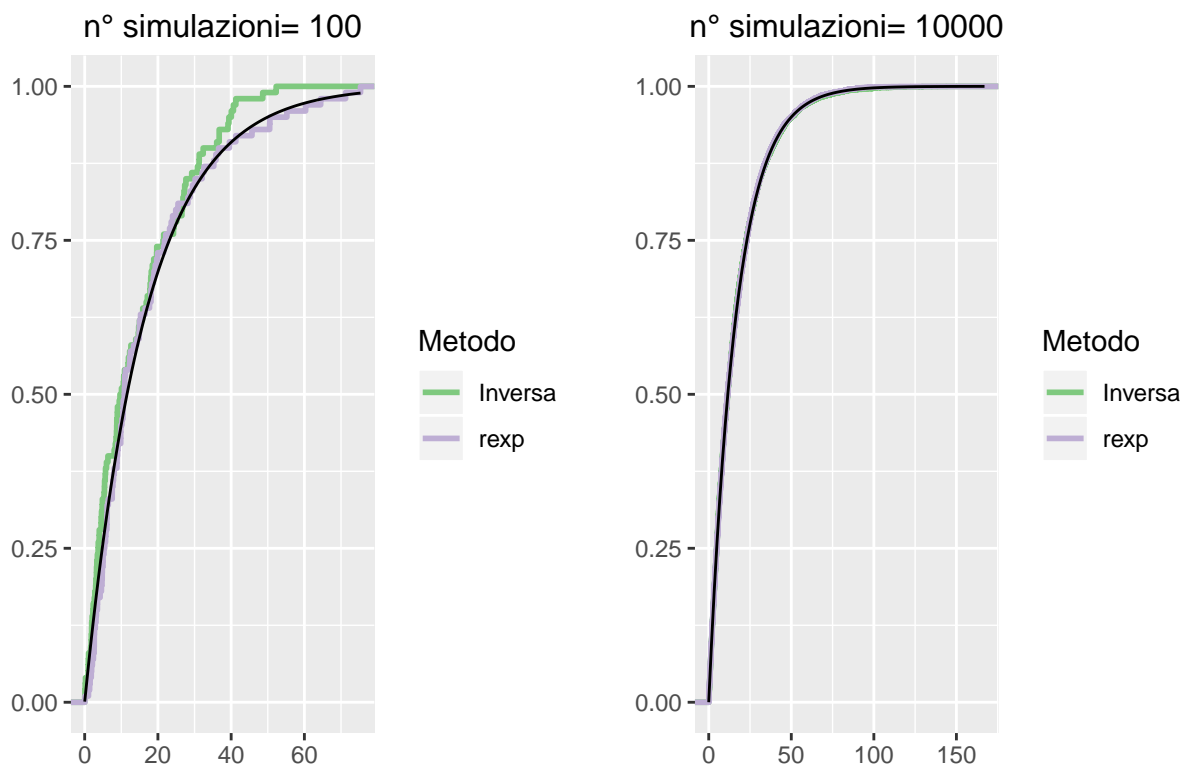
## Confronto densità empirica-teorica



### 1.4.3 Confronto dei due metodi

Infine si confrontano le funzioni di ripartizione empiriche, le medie e le varianze per entrambi gli algoritmi ed il numero di simulazioni. Si nota come con  $n = 10000$  i grafici si sovrappongono quasi completamente.

## Confronto funzione di ripartizione empirica-teorica

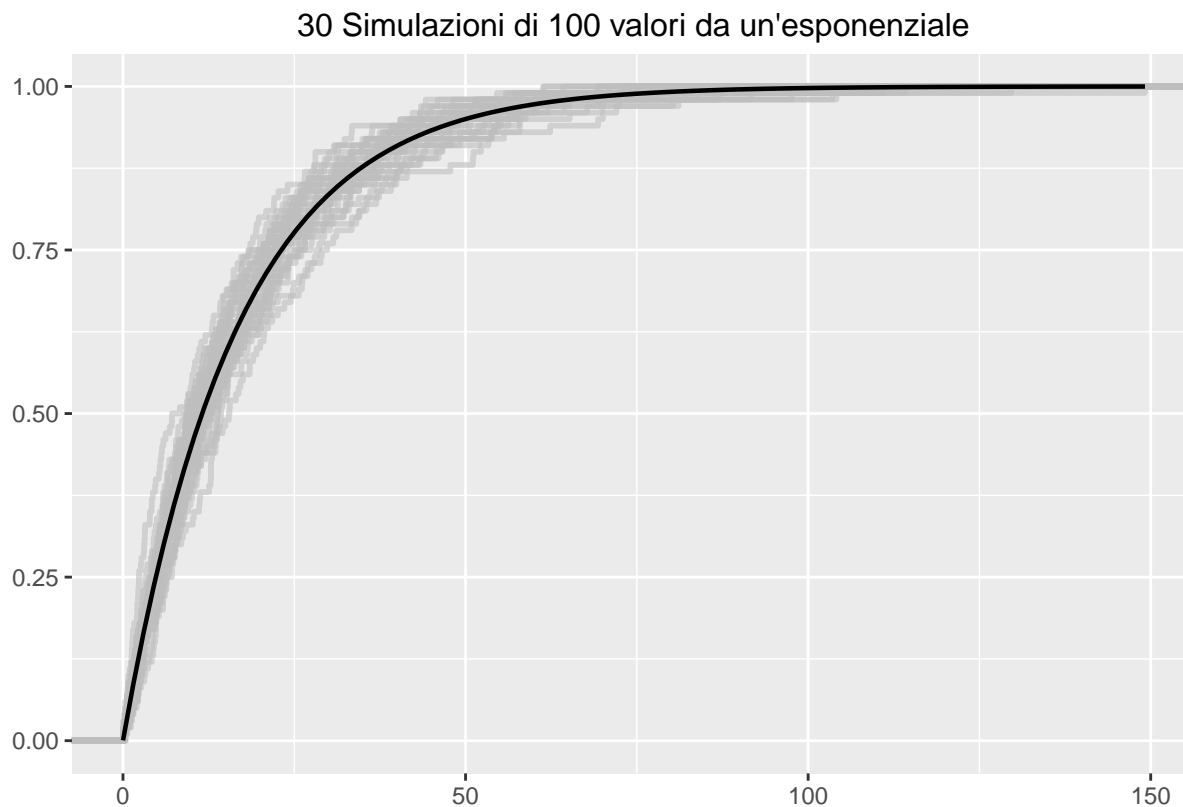


I momenti  $\mu$  e  $\sigma^2$  sono rispettivamente 16.7 e 277.8. Nella seguente tabella si mostrano i valori empirici.

	m-Inversa	var-Inversa	m-rexp	var-rexp
n=100	13.863	139.54	16.877	238.540
n=10000	16.460	266.28	16.691	273.269

#### 1.4.4 Simulazioni ripetute per n=100

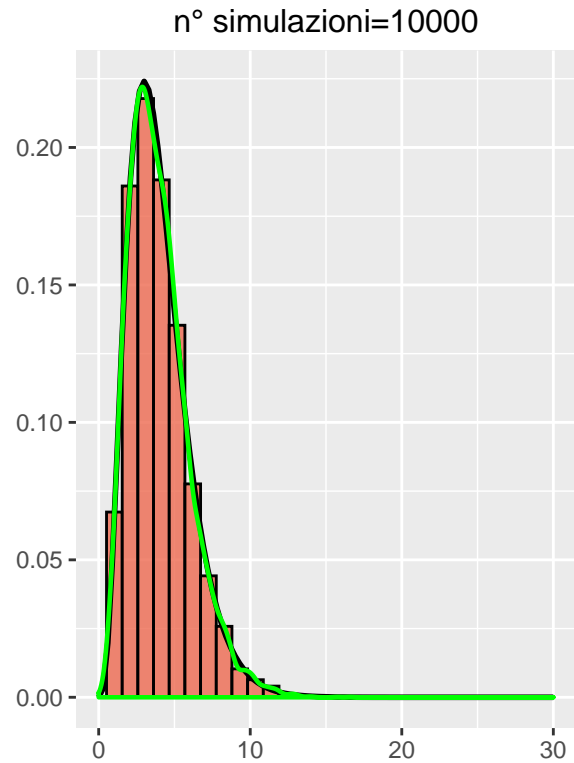
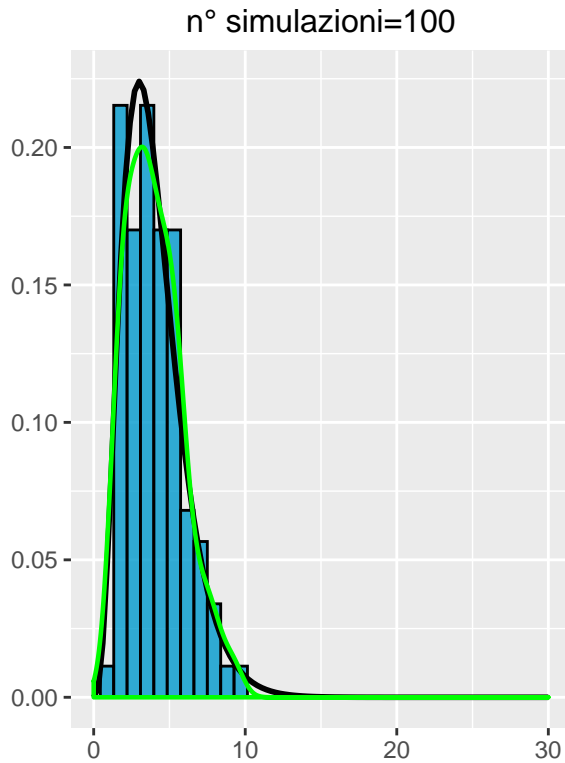
Si può anche verificare che, ripetendo le simulazioni per un fissato  $n$ , ad esempio  $n = 100$ , le funzioni di ripartizione empiriche hanno grafici che si avvicinano alla funzione di ripartizione empirica, alcune volte sovrastimando il vero valore, altre sottostimando.



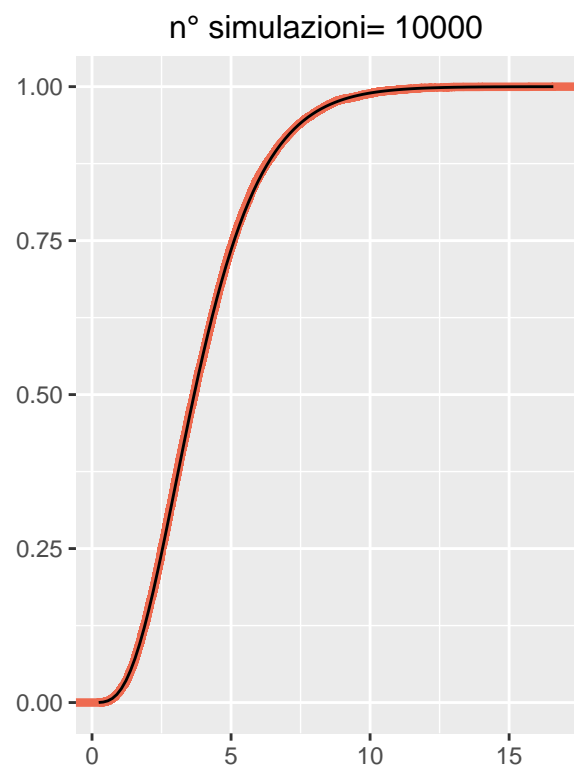
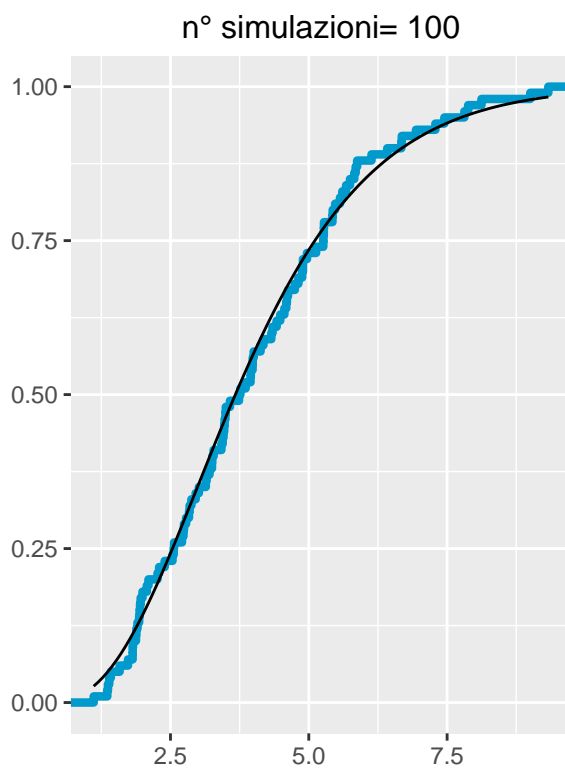
### 1.5 Distribuzione Gamma

Per simulare da una  $Ga(\alpha, \beta)$  si usa la funzione `rgamma(n, shape, rate)`.

## Confronto densità empirica-teorica



## Confronto funzione di ripartizione empirica-teorica



I momenti teorici sono  $\mu = \frac{\alpha}{\beta} = 4$  e  $\sigma^2 = \frac{\alpha}{\beta^2} = 4$ . Nella seguente tabella si mostrano i corrispondenti valori empirici.

	Media	Varianza
n=100	3.993	3.370
n=10000	3.988	3.981

## 1.6 Distribuzione di Poisson

### 1.6.1 Simulazione mediante l'algoritmo

Si simula da una  $Po(\lambda = 2)$ . L'algoritmo sfrutta i seguenti risultati:

$$\{N(t), t \geq 0\}, \quad N(t) \sim Po(\lambda t) \text{ e } N(1) \sim Po(\lambda)$$

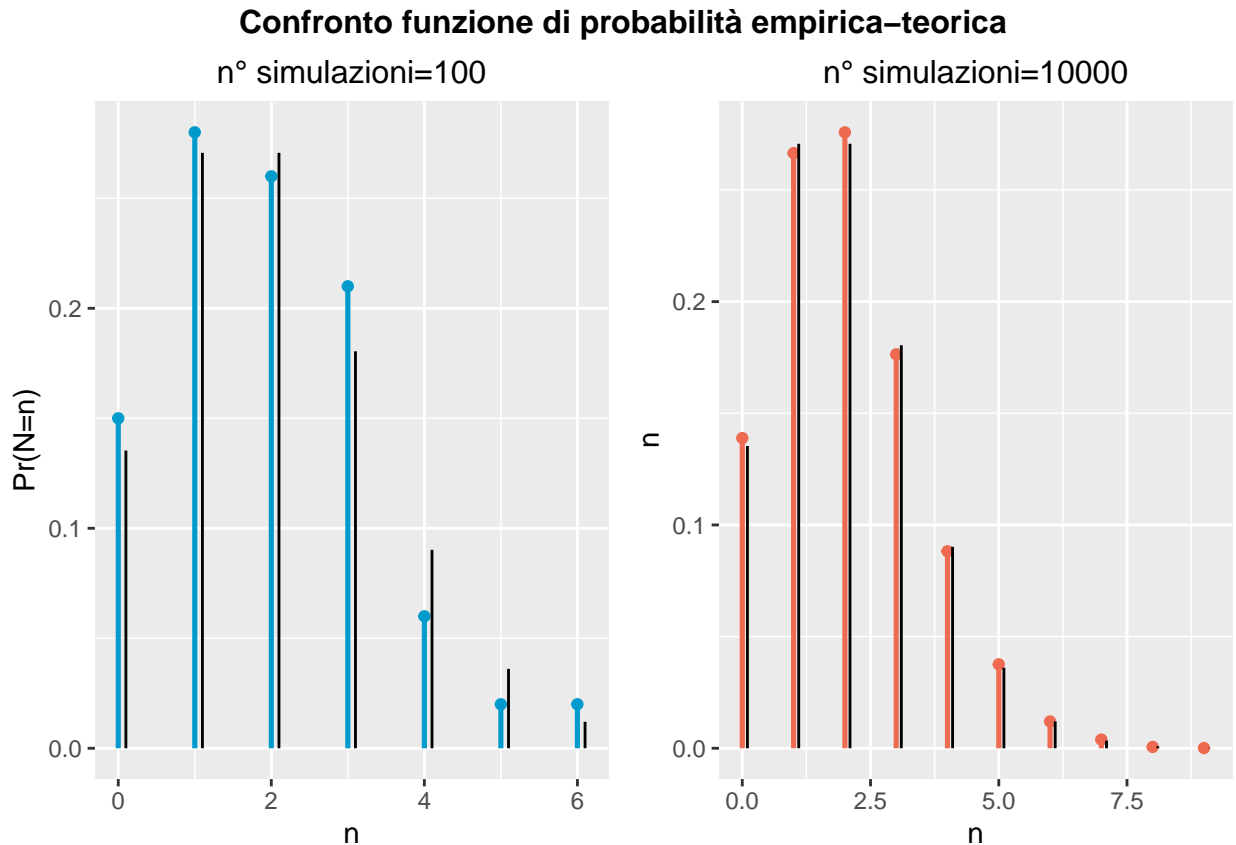
. Il numero aleatorio  $N(t)$  può essere riscritto nel seguente modo:

$$\sum_{n \geq 0} n | W_1 + \dots + W_n \leq t < W_1 + \dots + W_{n+1} |$$

dove i  $W_i$  sono i tempi di interarrivo ( $W_i$  è il tempo tra l'arrivo  $i - 1$  e l'arrivo  $i$ ), distribuiti come un' $Exp(\lambda)$ . Per prima cosa si simulano i tempi di interarrivo sequenzialmente (si simula da  $U \sim Unif(0, 1)$  ( $u'$  determinazione simulata), si calcola  $-\frac{1}{\lambda} \log(u')$  e si verifica il valore dell'indicatore d'evento nell'espressione sopra, con  $t = 1$ . Con semplici passaggi algebrici si prova che verificare l'espressione di destra in  $N = n \leftrightarrow -\frac{1}{\lambda} \log(U_1) + \dots + -\frac{1}{\lambda} \log(U_n) \leq 1 < -\frac{1}{\lambda} \log(U_1) + \dots + -\frac{1}{\lambda} \log(U_{n+1})$  equivale a verificare che

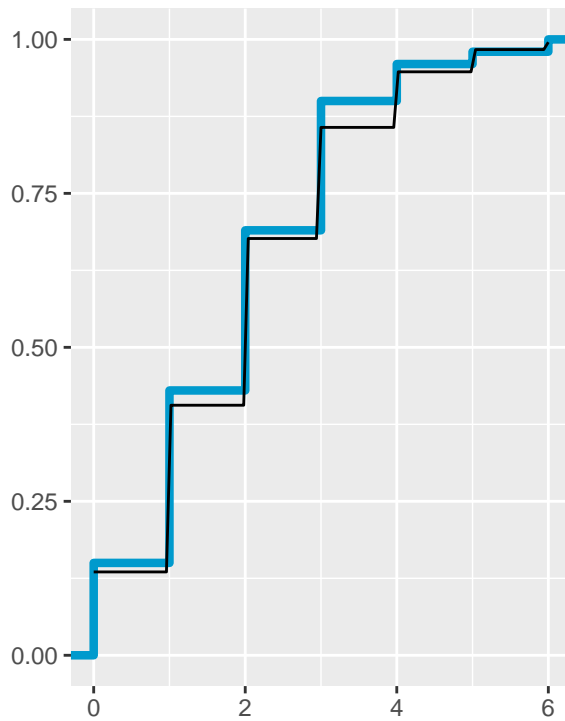
$$\prod_{i=1}^n U_i \geq -e^\lambda > \prod_{i=1}^{n+1} U_i$$

La funzione di probabilità teorica è rappresentata in nero.

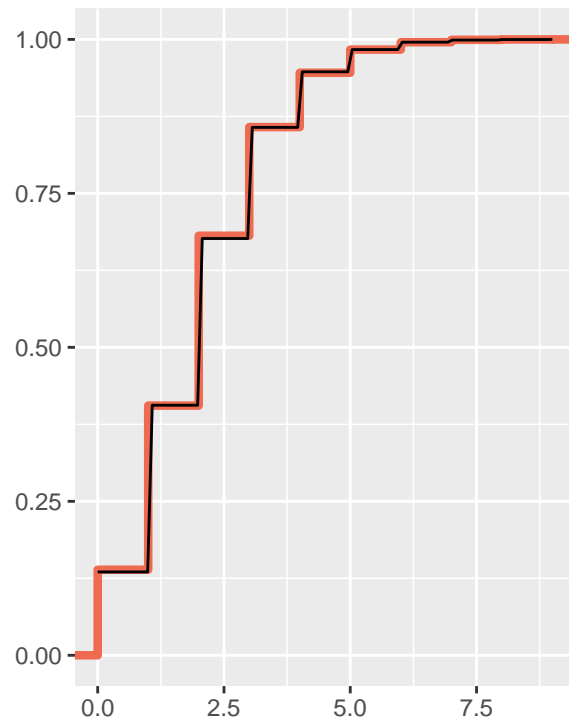


## Confronto funzione di ripartizione empirica–teorica

n° simulazioni= 100



n° simulazioni= 10000



### 1.6.2 Usando la funzione inversa (a gradini)

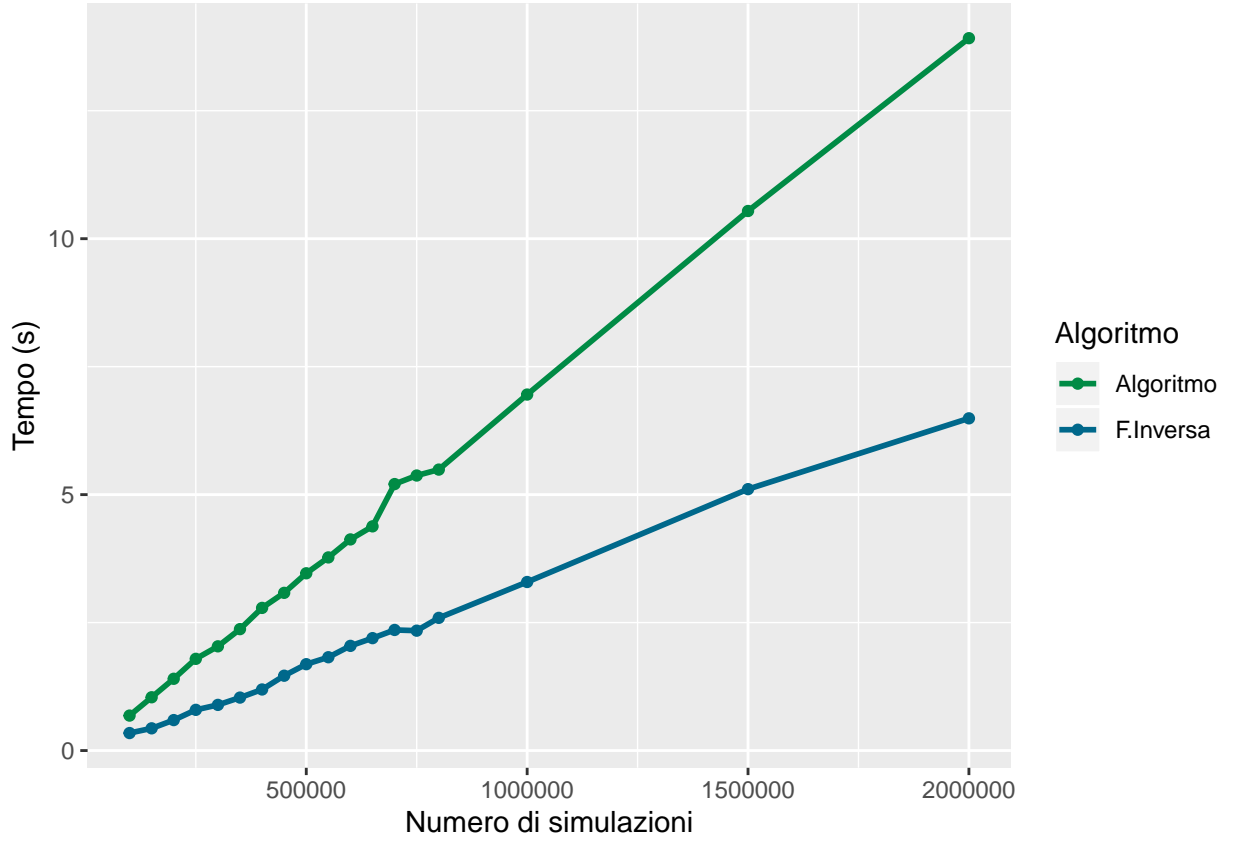
Un'alternativa per simulare da una  $X \sim Poiss(\lambda)$  si basa sulla  $F_X^{\leftarrow}(p)$ . Un risultato dice, infatti, che se  $X \sim F_X$  e  $U \sim Unif[0, 1]$ , la variabile aleatoria  $X' = F_X^{\leftarrow}(U)$  ha funzione di ripartizione  $F_X$ . Si conosce inoltre l'inversa generalizzata, in quanto per una funzione di ripartizione a gradini si ha che  $F_X^{\leftarrow}(p) = n_i \leftrightarrow F_X(n_{i-1}) < p \leq F_X(n_i)$ . L'idea è quindi di simulare da un'uniforme e verificare dove si colloca la determinazione simulata  $u'$  rispetto ai valori possibili della  $F_X$  in  $n_0, n_1, n_2, \dots$ .

### 1.6.3 Confronto dei due metodi

I momenti  $\mu$  e  $\sigma^2$  sono entrambi uguali a  $\lambda$  ( $= 2$ ). Nella seguente tabella si mostrano i valori empirici.

	m.Algoritmo	m.Inversa	var.Algoritmo	var.Inversa
n=100	1.890	2.040	1.836	1.635
n=10000	1.993	1.979	2.006	1.966

Inoltre, si può verificare il tempo che i due metodi impiegano per effettuare un certo numero  $n$  di simulazioni, al variare di  $n$ .



## 1.7 Distribuzione Poisson-composta

### 1.7.1 Metodo diretto

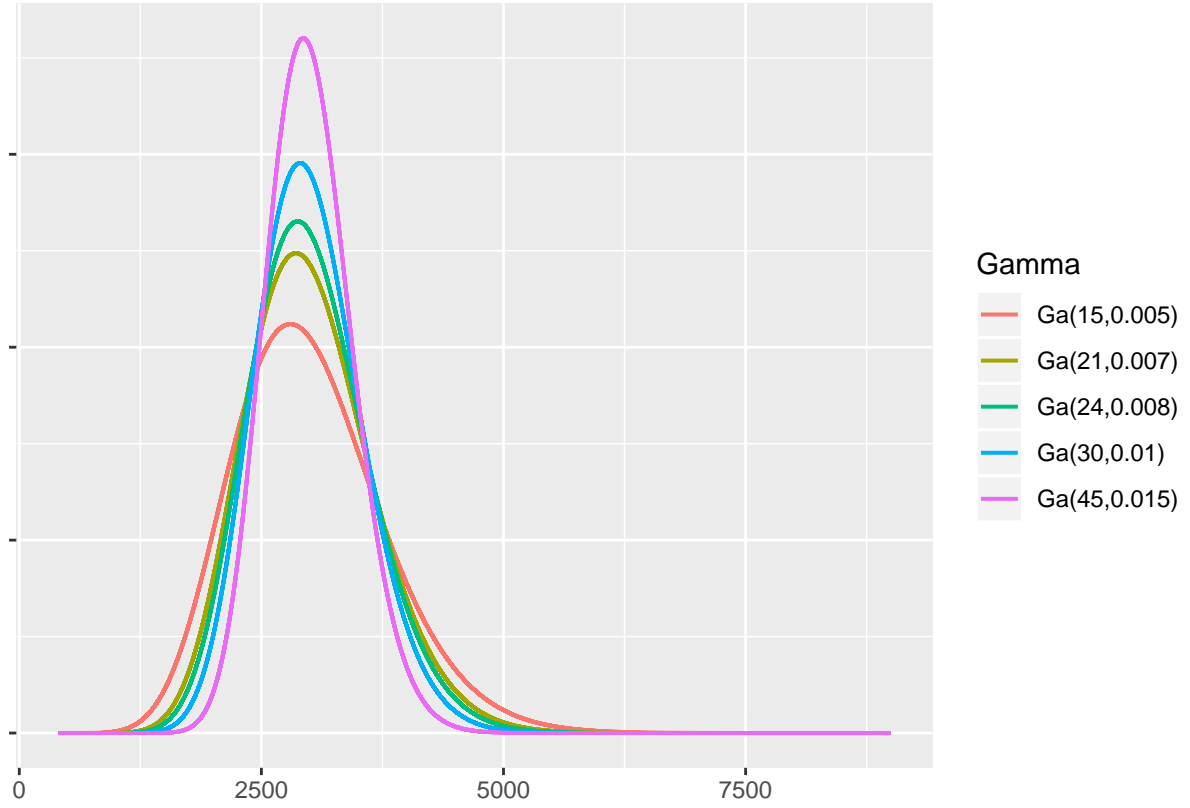
Si dice che  $X = \sum_{i=1}^N Y_i \sim PoissComposta(\lambda, F_Y)$  se vale che  $N \sim Poiss(\lambda)$  e  $Y_i|N = n, \dots, Y_n|N = n$  sono variabili aleatorie *I.I.D* con funzione di ripartizione  $F_y$  che non dipende da  $n$  (e neanche da  $i$ , con  $i \leq n$ , per ipotesi di identica distribuzione condizionata).  $F_Y$  è quindi la funzione di ripartizione comune, che rappresenta la funzione di ripartizione del sinistro nell'ipotesi che il sinistro si verifichi.  $X$  è quindi funzione del processo sottostante  $(\gamma(N, Y_1, Y_2, \dots))$ . La simulazione procede nel seguente modo:

- Si simula  $N \sim Poiss(\lambda)$  e si ottiene la distribuzione simulata  $n'$ ;
- Se  $n' = 0$  si pone  $x' = 0$ , dove  $x'$  è la determinazione simulata di  $X$ . Altrimenti si simulano  $n'$  valori dalla  $F_Y$ , ovvero si simulano dapprima  $n'$  uniformi  $u_1, \dots, u_{n'}$  e si calcolano poi:

- $y_1 = F_{Y_1|N=n'}^{\leftarrow}(u_1)$
- $F_{Y_2|N=n', Y_1=y_1}^{\leftarrow}(u_2) = F_{Y_2|N=n'}^{\leftarrow}(u_2)$  per l'ipotesi di indipendenza delle  $Y_i|N = n'$ ;
- ...

- la determinazione simulata di  $X$  è  $x' = \sum_{i=1}^{n'} y'_i$  e la determinazione simulata del processo sottostante è  $(n', y'_1, y'_2, \dots, y'_{n'}, 0, 0, 0, 0, \dots)$ ;

La sequenza va ripetuta un numero di volte pari al totale delle determinazioni simulate della  $X$  che si vuole avere. Nel seguente esempio si assume che la  $F_Y$  sia una  $Ga(\alpha, \rho)$  con parametri scelti in modo tale che la media  $\frac{\alpha}{\rho} \approx 3000$  e la varianza sufficientemente grande.



Si userà una  $Ga(15, 0.005)$  come  $F_Y$ , mentre per  $N$  si è scelto di usare  $\lambda = 60$ , ipotizzando un portafoglio di 1000 polizze con una frequenza sinistri pari a 0.06.

Di seguito si mostrano le densità e f.d.r empiriche, e un confronto tra medie e varianze campionarie e teoriche.

### 1.7.2 Approssimazione Wilson-Hilferty

L'approssimazione di Wilson Hilferty per simulare una  $X \sim PoissComposta(\lambda, F_Y)$  si basa su una trasformazione della  $X$  tramite una funzione  $h(\cdot)$  definita nel seguente modo:

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

, dove  $c_1, c_2$  e  $c_3$  sono coefficienti che dipendono da  $\mu_x, \sigma_x$  e dal coefficiente di asimmetria della  $X$ . L'approssimazione nasce dal fatto che vi è un risultato che afferma che

$$F_X(x) \approx \Phi\left(h\left(\frac{x - \mu_x}{\sigma_x}\right)\right)$$

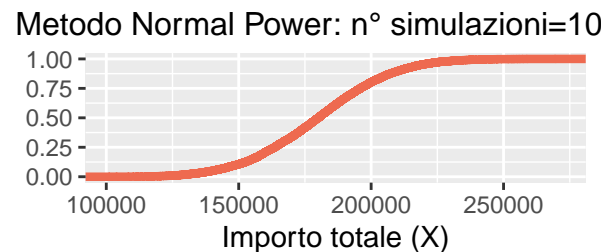
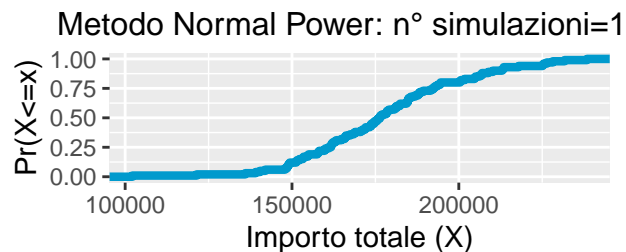
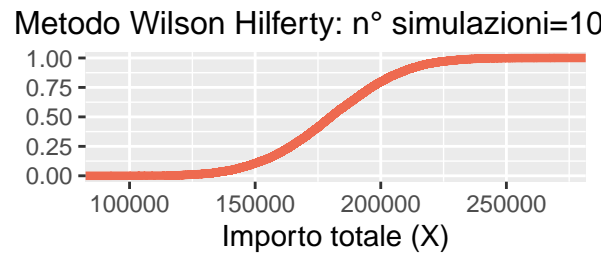
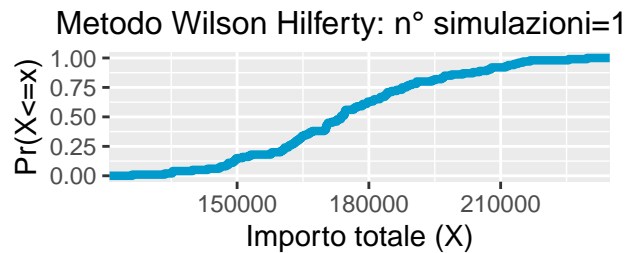
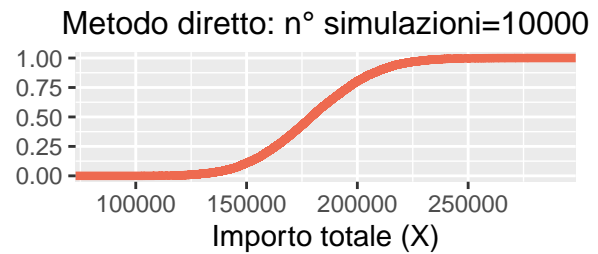
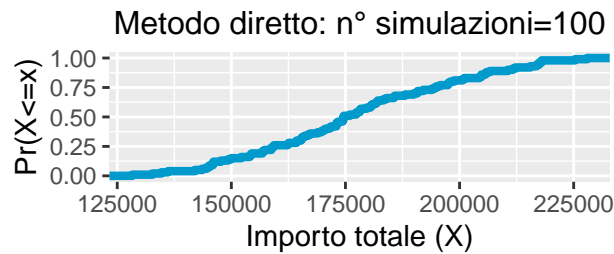
, ovvero, esplicitando, l'espressione di destra equivale a calcolare  $Pr(\sigma_x h^{-1}(W) + \mu_x \leq x)$ , con  $W \approx N(0, 1)$ . Il risultato permette di simulare quindi da  $X' = \sigma_x h^{-1}(W) + \mu_x$  e considerare i valori simulati come determinazioni simulate della  $X$  (in quanto  $F_{X'} \approx F_X$ ).

### 1.7.3 Approssimazione Normal Power

Anche l'approssimazione di *Normal Power* si basa sul calcolo di  $\mu_X, \sigma_X$  e  $\gamma_X$  per simulare da  $X' = \mu_X + \sigma_X W + \frac{\sigma_X \gamma_X}{6}(W^2 - 1)$ , con  $W \sim N(0, 1)$ . C'è un risultato, infatti, che dice che  $F_X \approx F_{X'}$ .

#### 1.7.4 Confronto metodo diretto, Wilson Hilferty e Normal Power

##### Funzioni di ripartizione empiriche



##### Confronto densità empiriche

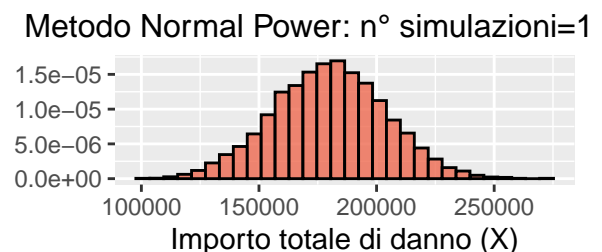
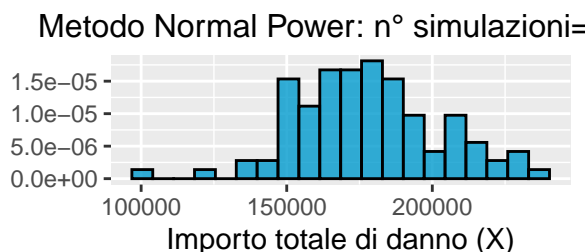
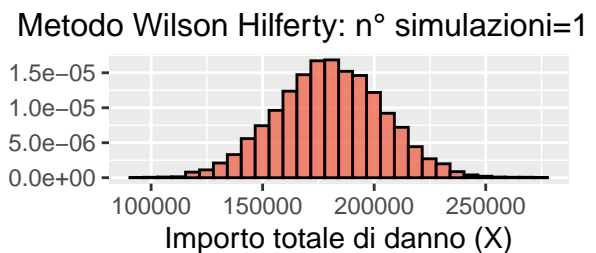
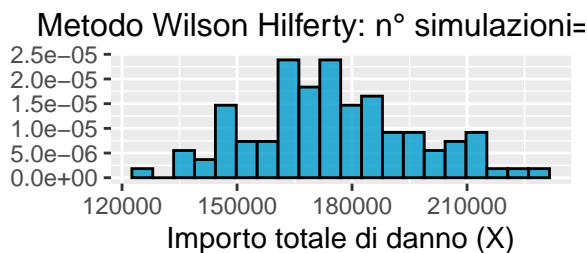
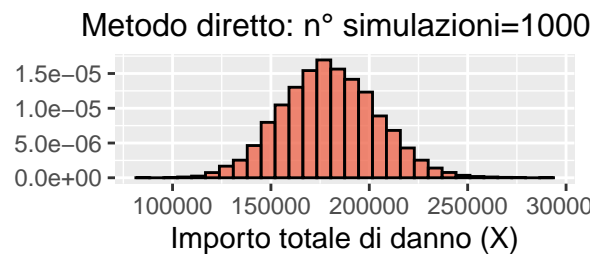
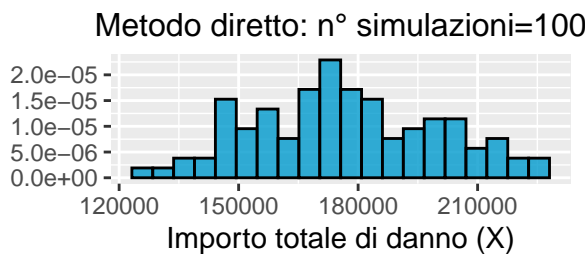


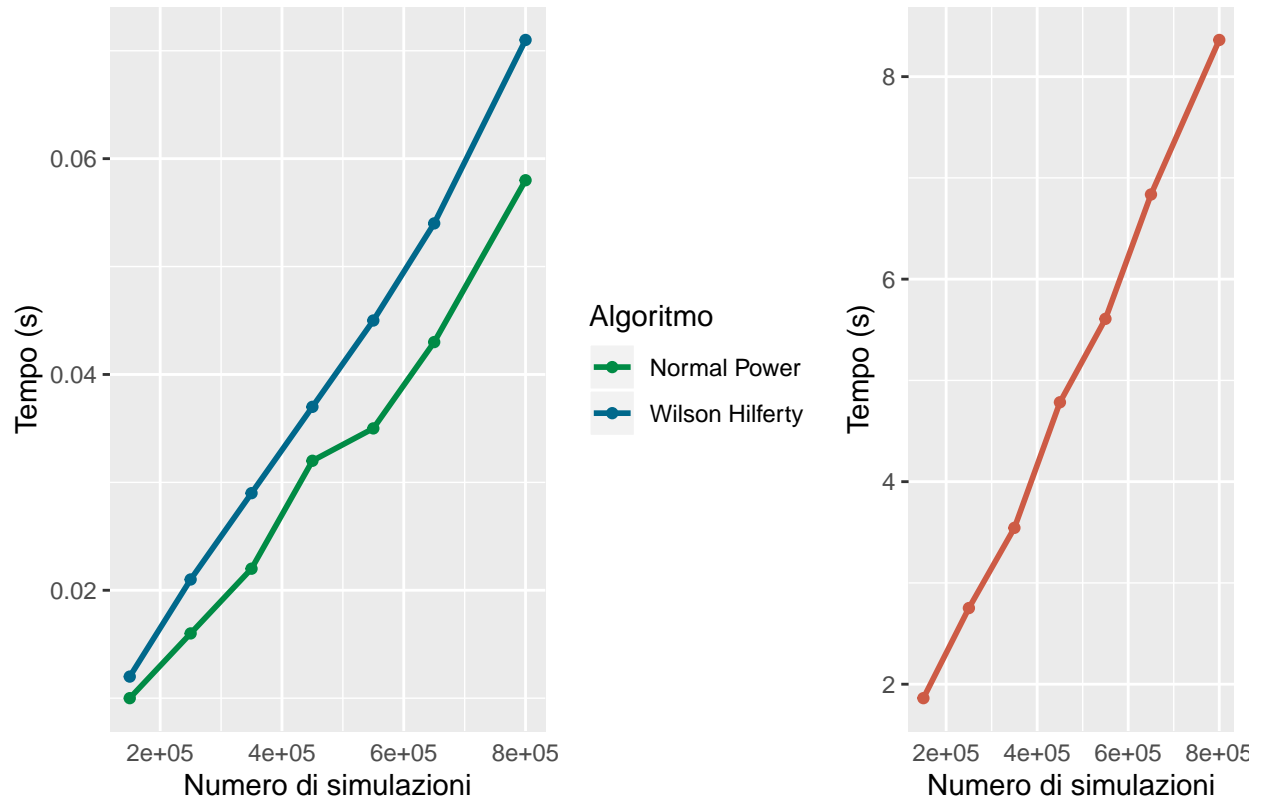


Table 1: Valori divisi per 1000

	M.Diretto n=100	M.Diretto n=10000	W.H. n=100	W.H. n=10000	N.P. n=100	N.P. n=10000	Teorica
Media	177.147	179.653	175.169	180.134	177.612	179.826	1.80e+02
Varianza	539115.116	580594.189	463.951	565.996	609.878	569.395	5.76e+08

Inoltre, si può verificare il tempo che i tre algoritmi impiegano per effettuare un certo numero  $n$  di simulazioni, al variare di  $n$ .

Confronto dei tempi per le simulazioni: si noti la diversa scala sull'asse delle y



## 2 Probabilità di rovina

**Consegna:** Valutare, mediante simulazione, la probabilità di rovina in un orizzonte limitato, nel modello della teoria collettiva del rischio.

Nel modello della teoria collettiva del rischio si ipotizza che il processo della rimanenza di portafoglio ( $R(t)$ ) sia definito nel seguente modo:

$$\{R(t), t \geq 0\} = \{R + P(t) - S(t) \mid t \geq 0\}$$

dove:

- $R$  è la dotazione di capitale iniziale di portafoglio;
- $P(t)$  è l'ammontare di premio incassato nell'intervallo  $[0, t]$ , e si ipotizza che sia proporzionale ad una costante  $c$  ( $P(t) = ct$ ) con  $ct$  maggiore del risarcimento atteso in  $[0, t]$ , ovvero  $c > \mu\lambda$  (con  $\mu$  valore atteso del singolo risarcimento e  $\lambda$  valore atteso del numero di sinistri in un intervallo unitario);
- $S(t)$  è l'ammontare totale pagato per sinistri avvenuti nell'intervallo  $[0, t]$  (si ipotizza che questi vengano pagati subito, e che non sia quindi riserva per sinistri avvenuti e non ancora risarciti).

Tale probabilità dipende da:

- $\tau$ : l'orizzonte temporale in cui si valuta la probabilità di rovina, che nel seguito sarà posto pari a 20;
- parametri del modello
  - $\lambda$ , parametro della Poisson;
  - $\alpha, \rho$  parametri della Gamma;
- caricamento del premio;
- $R$ , capitale inizialmente allocato al portafoglio.

L'evento rovina nell'orizzonte  $[0, \tau]$  si verifica se  $\exists t \in [0, \tau] : R(t) < 0$ . Per valutare la probabilità dell'evento rovina si simulano un numero  $n$  di traiettorie e si calcola la frequenza con cui l'evento rovina si presenta.

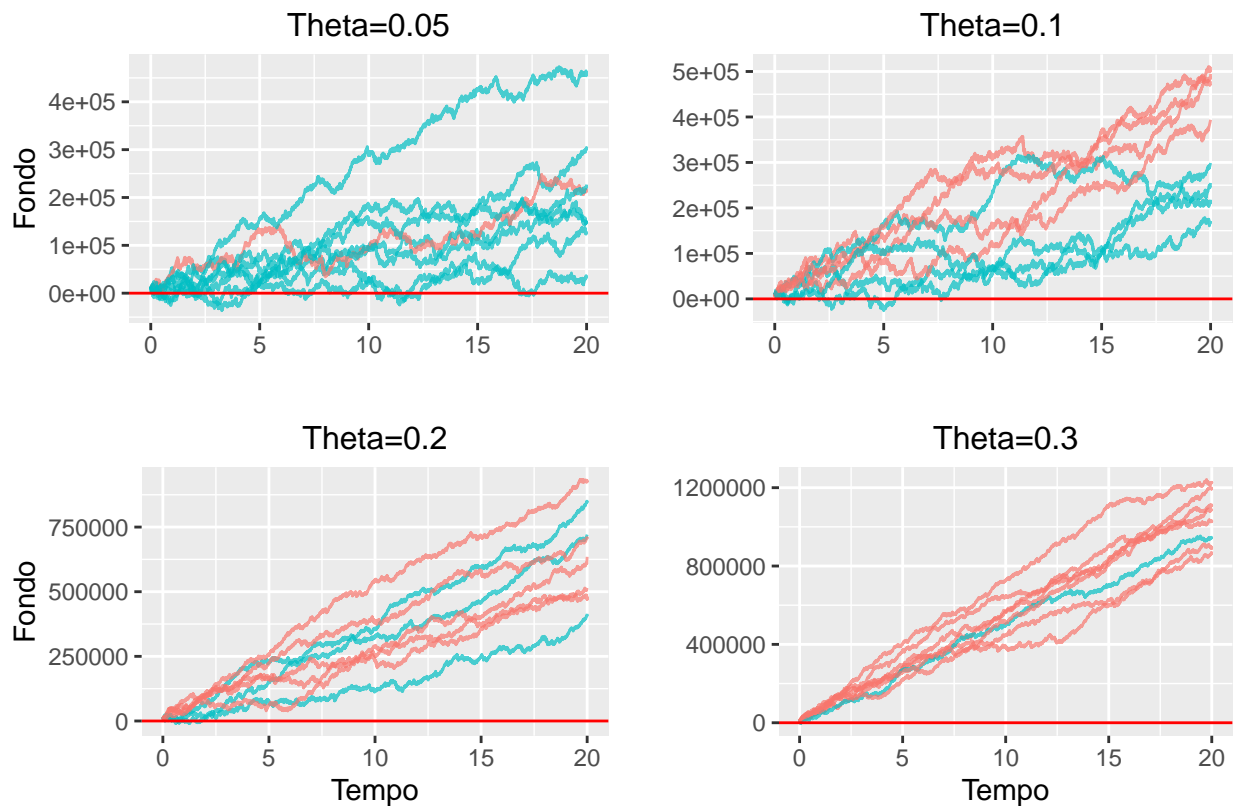
Si mostrano innanzitutto 8 traiettorie simulate al variare dei parametri da cui le traiettorie (e quindi la probabilità di rovina) dipendono.

Si fissano inizialmente le seguenti variabili: si ipotizza di avere un fondo iniziale di 10.000 euro e si immagina di

avere un portafoglio costituito da 1000 polizze. Ipotizzando una frequenza sinistri per polizza di 0.06, si è scelto di usare  $\lambda = 60$ . Si utilizza come  $F_Y$  una  $Ga(15, 0.005)$ , che ha quindi una media pari a 3000 e presenta varianza abbastanza elevata.

### 2.1 Probabilità di rovina al variare di Theta

I primi grafici mostrano come variano le traiettorie al variare del **coefficiente di caricamento del premio Theta**.

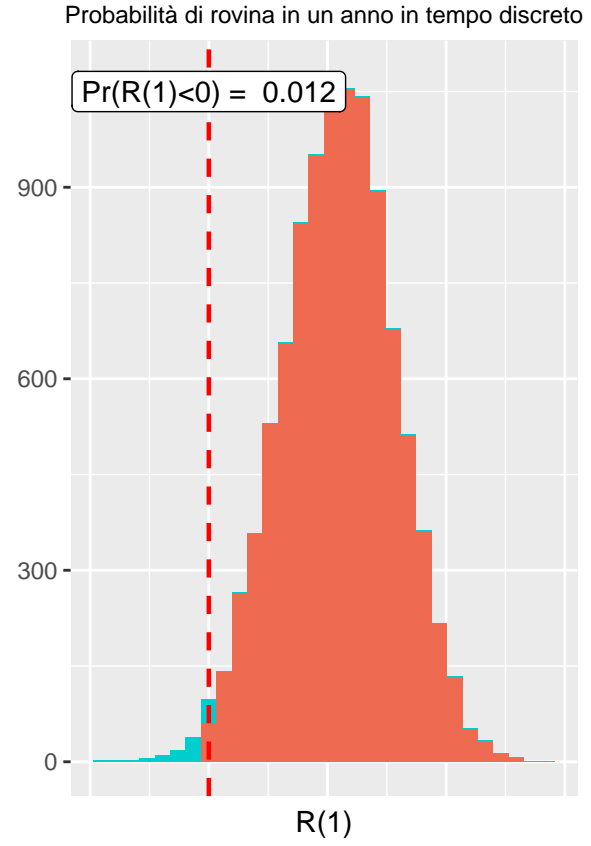
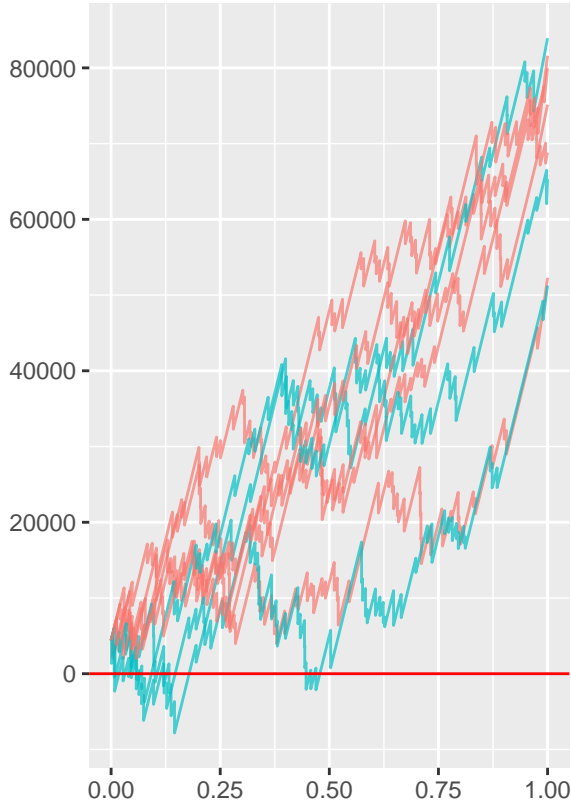


Probabilità di rovina al variare di Theta: in azzurro le traiettorie che portano alla rovina

Si nota che la probabilità dell'evento rovina decresce al crescere del caricamento  $\Theta$ . Si vede, inoltre, che spesso l'evento rovina si verifica per  $t$  molto vicini a 0, a causa della variabilità iniziale. Per diminuire la probabilità di avere nei primi anni  $R(t) < 0$  una soluzione è allocare più capitale iniziale  $R$ .

Per esempio, si prova ad allocare un  $R$  pari alla differenza tra il quantile empirico di ordine 0.99 delle determinazioni simulate della  $X = \sum_{h=1}^N Y_h$  (dove  $X$  ha distribuzione composta  $(\lambda, F_y)$ ) e l'ammontare di premi incassati in un anno (che secondo le ipotesi sono  $\mu\lambda(1 + \Theta)$  con  $\Theta = 0.3$ ,  $\mu = 3000$  e  $\lambda = 60$ ). Facendo questo ragionamento, risulta che il capitale da allocare è pari a 4350.4063142.

Si nota, però, che questo limita la probabilità di rovina in tempo discreto nell'orizzonte limitato di un anno ( $Pr(R(1) < 0)$ ), ma non è sufficiente per limitare la probabilità di rovina per  $t$  molto vicini a 0, come si vede dalle traiettorie sotto rappresentate.



## 2.2 Probabilità di rovina al variare di $F_Y$

Si fissa ora  $\theta = 0.3$  e si simulano traiettorie dove i parametri  $\alpha$ ,  $\rho$  della  $F_Y$  variano:

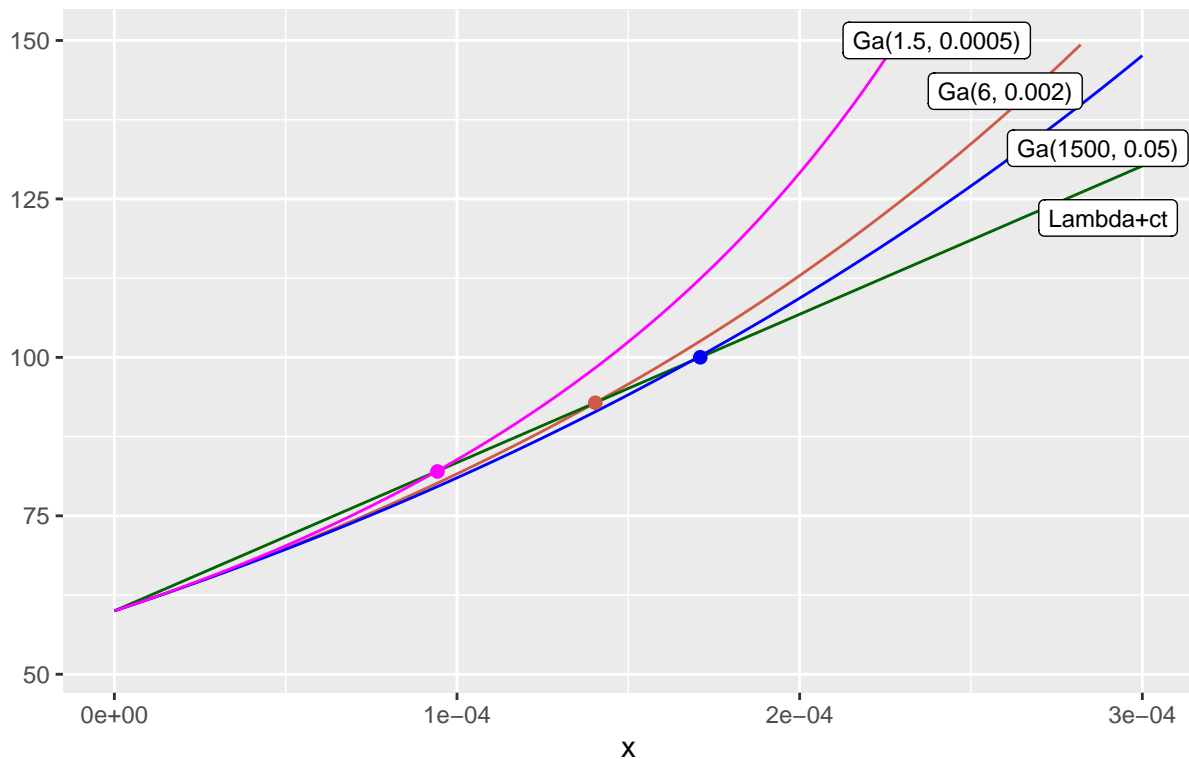
- La media resta fissa e pari a 3000;
- La varianza aumenta al diminuire di  $\rho$ , in quanto la varianza è  $\frac{\alpha}{\rho} \frac{1}{\rho} = \frac{3000}{\rho}$

Di seguito sono mostrate le due densità Gamma utilizzate e le corrispondenti traiettorie e probabilità di rovina. Si nota che la probabilità di rovina diminuisce al diminuire di  $\sigma_Y^2$ . Questo risultato si può giustificare ricorrendo al seguente teorema: sotto le ipotesi della teoria collettiva del rischio precedentemente enunciate (e alcune ipotesi aggiuntive),  $\exists! \alpha > 0$ , ( $\alpha \in [0, \gamma[$ ) soluzione di  $\lambda m_Y(t) = \lambda + ct$ . La soluzione dell'equazione è importante perchè permette di limitare la probabilità di rovina asintotica, infatti vale che

$$\Psi(R) \leq e^{-\alpha R}$$

Tenendo fissi tutti i parametri tranne  $\alpha$ ,  $\rho$ , ma mantenendo costante il loro rapporto, ciò che si modifica è solo la  $m_Y(t) = (\frac{\rho}{\rho-t})^\alpha$ . Vediamo graficamente come si muove la soluzione dell'equazione quando la varianza diminuisce, ma  $\frac{\alpha}{\rho}$  resta costante.

### Alpha al variare dei parametri della Gamma

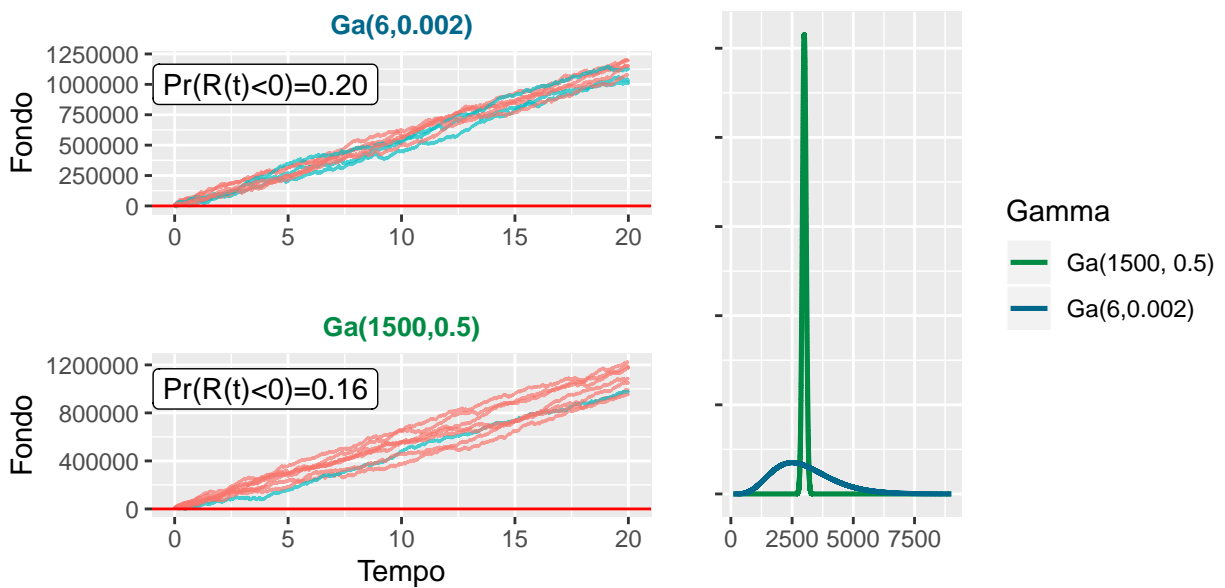


La soluzione dell'equazione è maggiore quando la  $F_y$  ha maggior varianza

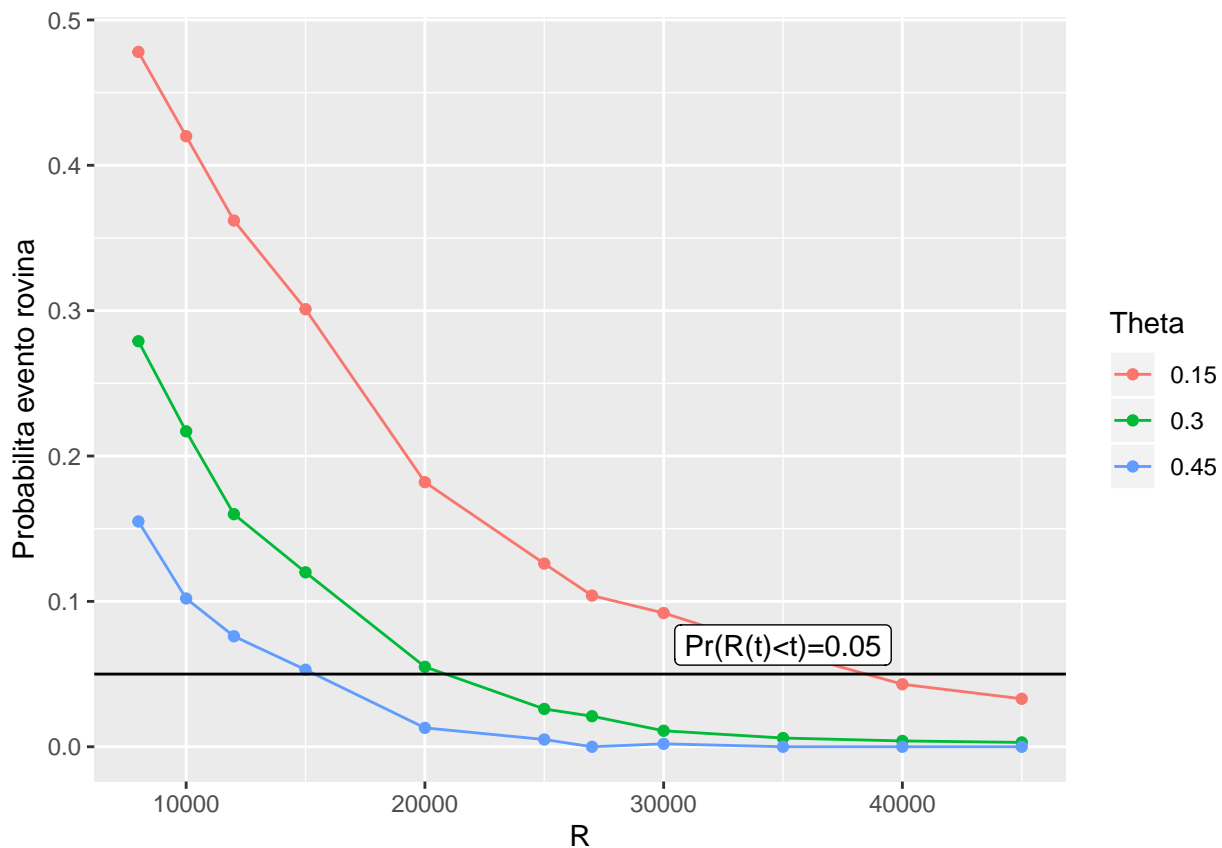
Da questo esempio si vede come, al diminuire della varianza ( $\rho$  più grande)  $\alpha$ , soluzione dell'equazione, cresce, quindi  $e^{-\alpha}$  decresce, ovvero la limitazione superiore della probabilità di rovina decresce.

Traiettorie al variare di Alpha, Rho: traiettoria rovina in azzurro.

Le probabilità di rovina solo calcolate su 5000 traiettorie simulate

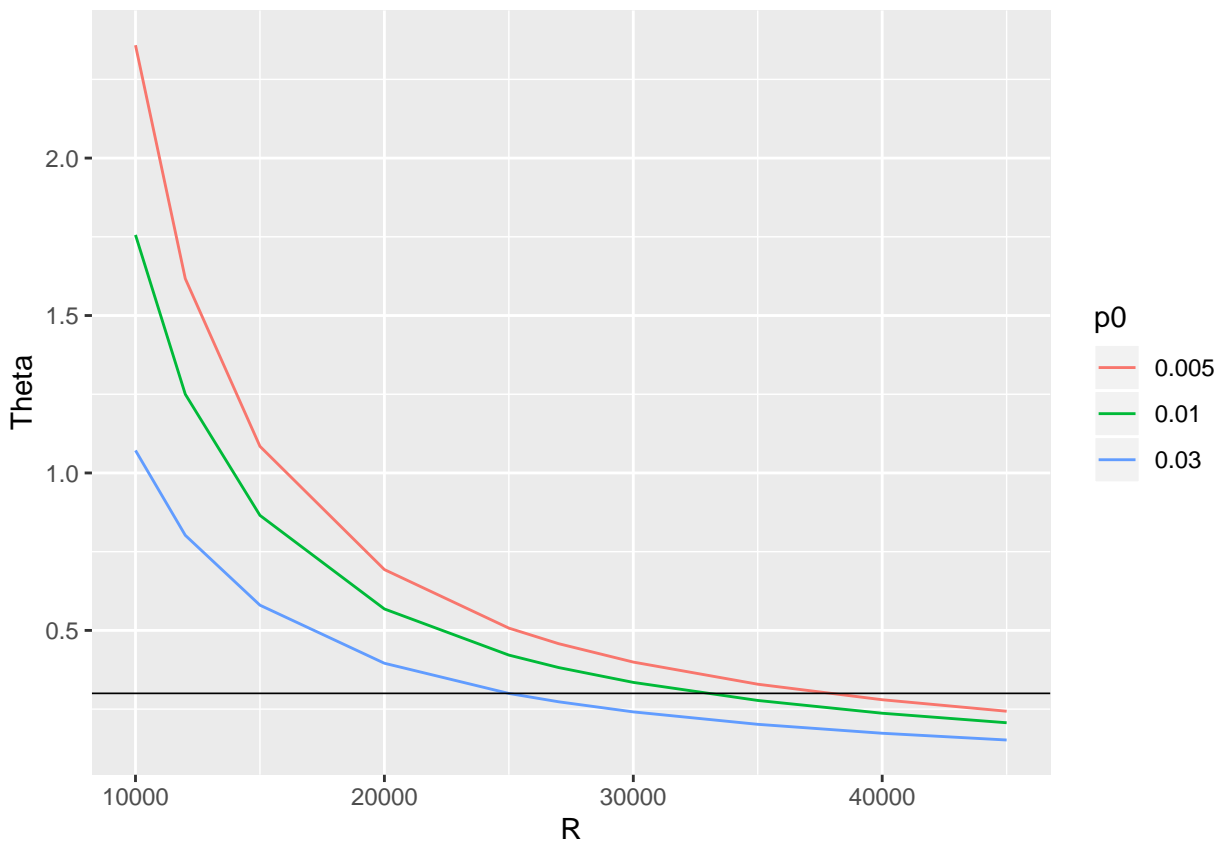


### 2.3 Probabilità di rovina al variare di Theta e R



Un altro modo per vedere quanto appena detto è considerare l'equazione  $\lambda m_Y(t) = \lambda + ct$ , fissare una data probabilità di rovina  $p_0$  e ricavare  $\theta$  in funzione di  $R$ . Il risultato è

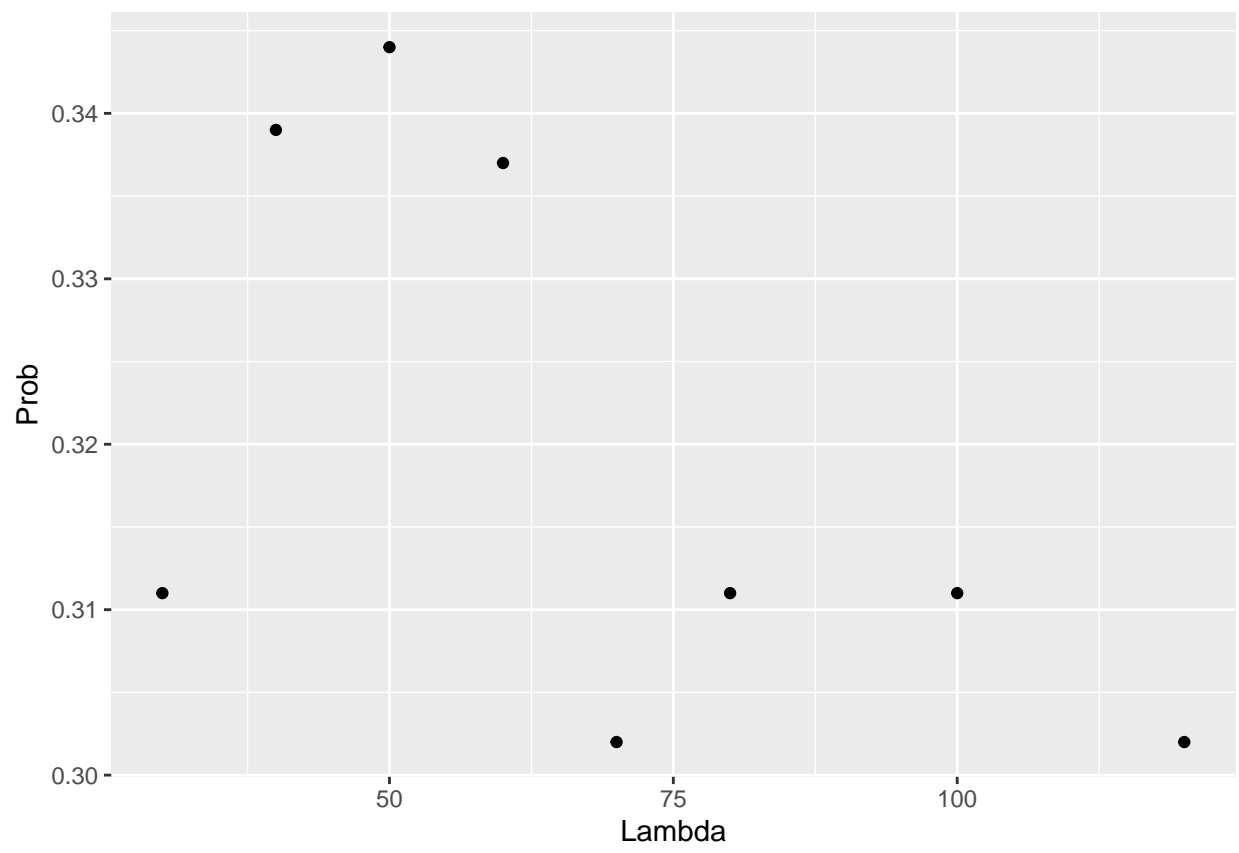
$$\theta = \frac{m_Y(-\frac{1}{R}\ln(p_0)) - 1}{\mu(-\frac{1}{R}\ln(p_0))} - 1$$



## 2.4 Probabilità di rovina al variare di Lambda

Si è detto che  $\Psi(R) \leq e^{-\alpha R}$  e che  $\alpha$  è la soluzione di  $\lambda m_Y(t) = \lambda + ct$ . Esplicitando  $c = (1 + \theta)\mu\lambda$  e dividendo per  $\lambda$  si vede che la soluzione in realtà è indipendente da  $\lambda$ .

Si verifica che la probabilità di rovina non cambia se si tengono fissi tutti i parametri e si varia solo  $\lambda$ .





### 3 Metodi deterministici per il calcolo delle riserve sinistri

Dati i seguenti dati, calcolare le riserve sinistri con i metodi deterministici presentati durante le lezioni.

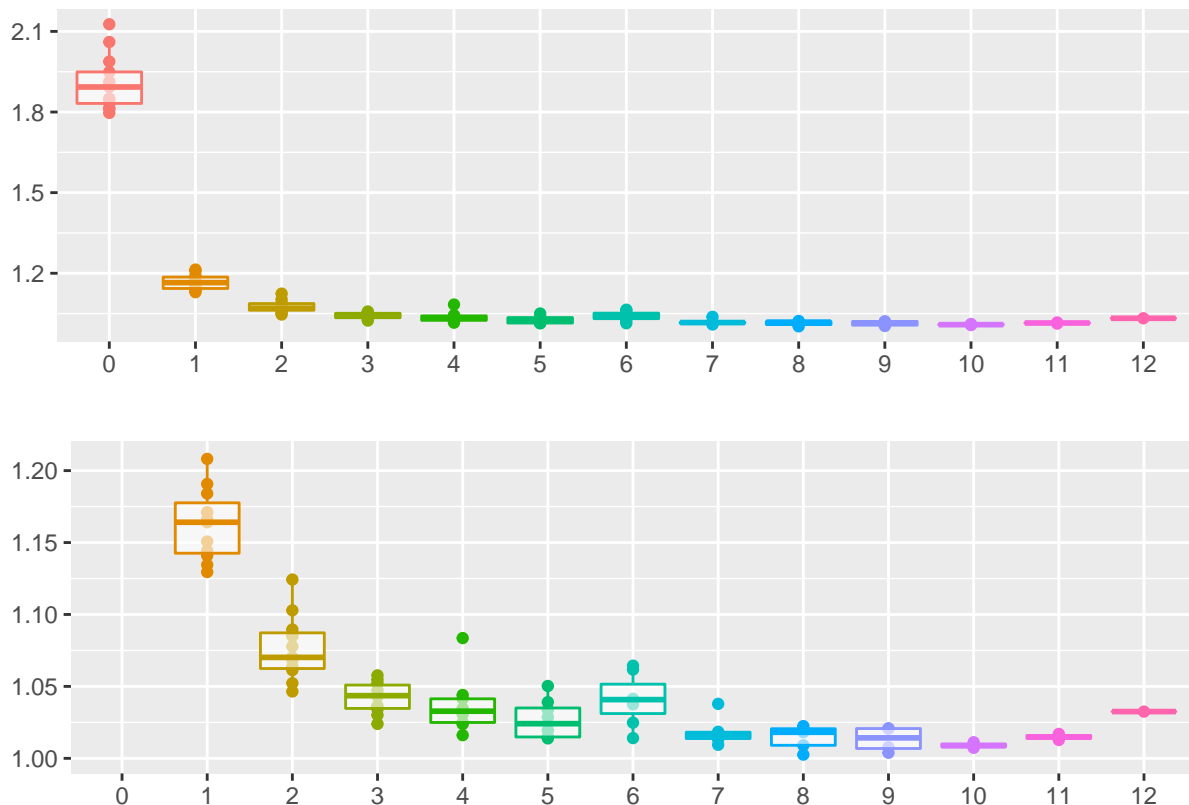
	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	10775462	19934512	22516307	23896590	24768314	25358202	26633758	28277294	28799583	29444853	30058889	30395282	30783352	31783405
1	11884473	21949545	25042924	27285391	28777663	29773540	30937259	32930435	34177549	34797083	35515587	35826518	36434605	
2	12758868	24869817	28951246	32549136	34089494	36938833	37995255	39545986	40064331	40888090	41046113	41345861		
3	13419185	24110891	28117970	30505355	32265113	33256169	33906857	35177601	35783677	36107762	36390079			
4	14484622	26271424	30231129	33341027	34806484	35946997	36473076	37982260	38580218	38678589				
5	16217856	32238045	37753900	40153294	41885972	43619021	44223116	44844786	45266549					
6	19363970	35474849	42005893	43957333	46248842	46992831	47694230	48875526						
7	19613399	35546749	40328006	42434459	43881583	45814679	47356681							
8	21210290	40154088	47812880	51087329	52626036	53937341								
9	22725871	43242432	49479958	52952532	54219997									
10	21153719	44987513	54352092	58581114										
11	21740344	41586121	50477110											
12	19870614	40946719												
13	21950030													

#### 3.1 Chain ladder

Il modello è:

$$f_{j-1} = \frac{C_{i,j}}{C_{i,j-1}}$$

Bisogna dunque verificare che il rapporto tra i pagamenti cumulati con differimento  $j$  e quelli con differimento  $j - 1$  siano più o meno costanti, al variare dell'anno di origine e per ogni  $j$  per vedere se il modello è applicabile.



La stima per  $f_j$ ,  $j \in \{0, \dots, t-1\}$  avviene nel seguente modo:

$$f_{j-1} = \frac{\sum_{i=0}^{t-j} C_{i,j}}{\sum_{i=0}^{t-j} C_{i,j-1}}$$

I valori stimati sono riportati in tabella:

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	10775462	19934512	22516307	23896590	24768314	25358202	26633758	28277294	28799583	29444853	30058889	30395282	30783352	31783405
1	11884473	21949545	25042924	27285391	28777663	29773540	30937259	32930435	34177549	34797083	35515587	35826518	36434605	37618249
2	12758868	24869817	28951246	32549136	34089494	36938833	37995255	39545986	40064331	40888090	41046113	41345861	41967816	43331216
3	13419185	24110891	28117970	30505355	32265113	33256169	33906857	35177601	35783677	36107762	36390079	36713319	37265588	38476228
4	14484622	26271424	30231129	33341027	34806484	35946997	36473076	37982260	38580218	38678589	39164100	39511980	40106349	41409276
5	16217856	32238045	37753900	40153294	41885972	43619021	44223116	44844786	45266549	45907258	46483506	46896402	47601852	49148285
6	19363970	35474849	42005893	43957333	46248842	46992831	47694230	48875526	49749900	50454066	51087388	51541179	52316500	54016096
7	19613399	35546749	40328006	42434459	43881583	45814679	47356681	49151005	50030307	50738443	51375334	51831683	52611373	54320549
8	21210290	40154088	47812880	51087329	52626036	53937341	55299807	57395093	58421880	59248791	59992509	60525400	61435868	63431723
9	22725871	43242432	49479958	52952532	54219997	56183346	57602547	59785082	60854626	61715970	62490656	63045738	63994119	66073084
10	21153719	44987513	54352092	58581114	60967819	63175512	64771336	67225494	68428145	69396686	70267784	70891947	71958356	74296054
11	21740344	41586121	50477110	54220290	56429327	58472677	59949707	62221175	63334300	64230742	65036995	65614695	66601720	68765397
12	19870614	40946719	47943452	51498746	53596902	55537688	56940580	59098033	60155286	61006732	61772516	62321218	63258700	65313774
13	21950030	42036154	49219043	52868930	55022910	57015333	58455550	60670405	61755787	62629887	63416045	63979347	64941771	67051522

Figure 1: Triangolo di Runoff: le stime sono evidenziate in giallo

j=0	j=1	j=2	j=3	j=4	j=5	j=6	j=7	j=8	j=9	j=10	j=11	j=12
1.915084	1.170874	1.074156	1.040742	1.036211	1.02526	1.03789	1.01789	1.014154	1.012552	1.008883	1.015043	1.032487

Infine si proiettano i pagamenti cumulati  $C_{i,j}$  con  $i + j > t$ . Ad esempio,  $\hat{C}_{4,10} = C_{4,9} \cdot \hat{f}_9$ , con  $C_{4,9}$  noto. Nella seguente tabella si vedono i dati proiettati nel triangolo inferiore.

Table 1: Stime dei parametri del modello

h	lambda_h	r_h
0	1252.562	0.3834344
1	1318.048	0.3260903
2	1318.873	0.1077417
3	1489.767	0.0498143
4	1612.995	0.0268099
5	1821.479	0.0226706
6	2163.910	0.0150882
7	2419.766	0.0216684
8	2458.260	0.0099073
9	2527.640	0.0072597
10	2714.751	0.0062985
11	3009.087	0.0042103
12	3116.694	0.0065974
13	3591.951	0.0124093

Figure 2: Stima dei futuri tassi di inflazione

### 3.2 Metodo di separazione (o di Taylor)

Il modello per i  $P_{i,j}$  è:

$$P_{i,j} = \lambda_{i+j} n_i r_j$$

con  $n_i$  il numero di sinistri totali denunciati nell'anno  $i$  (noti). Sia

$$s_{i,j} = \frac{P_{i,j}}{n_i}$$

. I  $\lambda_h$  e gli  $r_j$  vengono stimati nel seguente modo:

$$\hat{\lambda}_h = \frac{\sum_{i+j=h} s_{i,j}}{1 - \sum_{j=h+1}^t \hat{r}_j} \quad e \quad \hat{r}_h = \frac{\sum_{i=0}^{t-h} s_{i,h}}{\sum_{j=h}^t \hat{\lambda}_{t,j}}$$

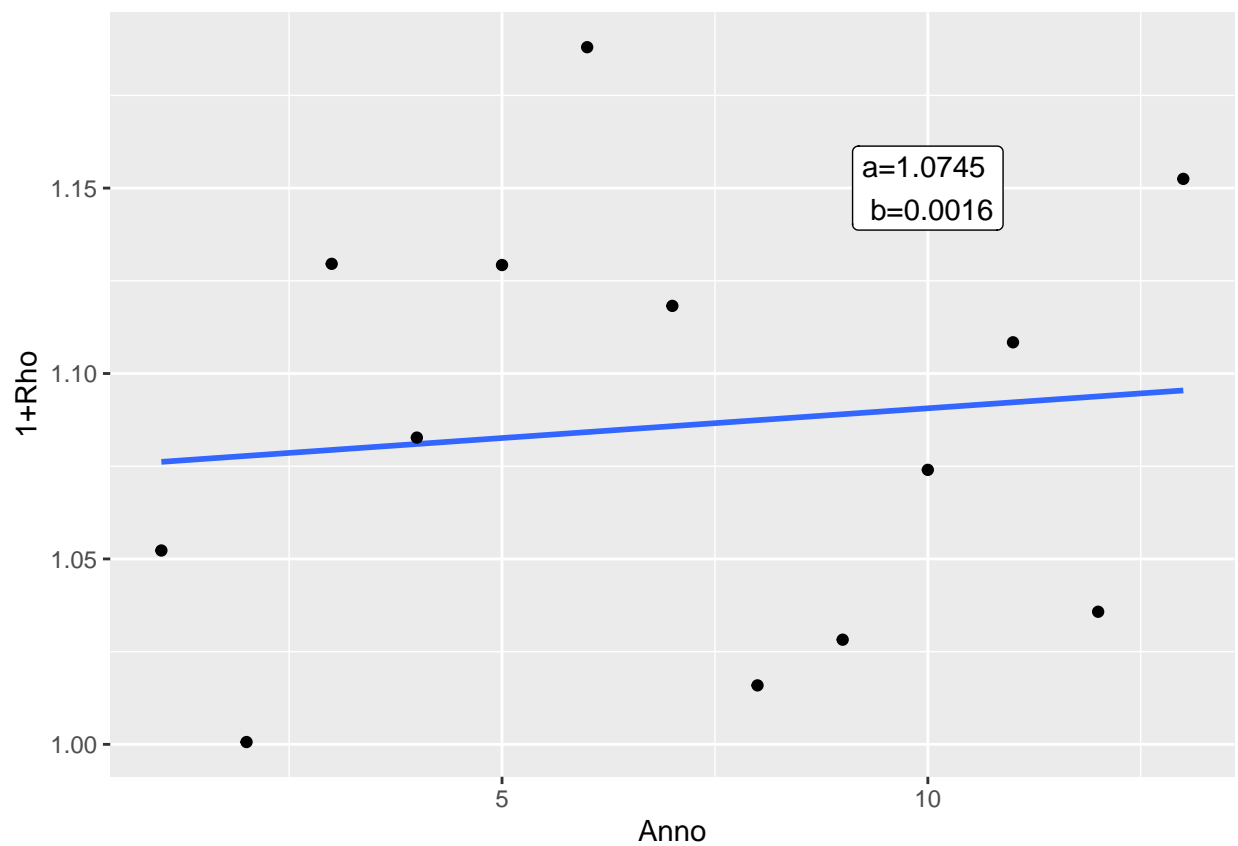
È poi necessario proiettare i  $\lambda_{t+1}$ ,  $\lambda_{t+2}$ , .... Una delle ipotesi del modello è che il costo per sinistro in termini reali  $\mu$  sia costante al variare dell'anno di origine, e  $\lambda_{i+j}$  prende quindi il significato di costo per sinistro espresso a valori correnti dell'anno  $i+j$  ( $\lambda_{i+j} = \mu(1+\rho_1)(1+\rho_2)\dots(1+\rho_{i+j})$ , con  $\rho_i$  coefficiente di adeguamento dei costi dall'anno  $i-1$  all'anno  $i$ ). Data questa osservazione, si può verificare che

$$\frac{\lambda_k}{\lambda_{k-1}} = 1 + \rho_k$$

Per stimare i futuri  $\lambda_k$  si estrapolano i tassi annui di inflazione futura dalle stime ottenute dal triangolo. Si cerca, quindi, una funzione che perequi i valori  $\frac{\hat{\lambda}_k}{\hat{\lambda}_{k-1}}$ . Per esempio, il problema può essere impostato in questo modo:

$$\min_{a,b} \sum_{k=1}^t \left( \frac{\hat{\lambda}_k}{\hat{\lambda}_{k-1}} - (a + bk) \right)^2$$

Si stima poi  $\hat{\lambda}_{t+1} = (\hat{a} + \hat{b}(t+1))\hat{\lambda}_t$ , e così i successivi.



Infine si proiettano i  $\hat{P}_{ij} = n_i \hat{\lambda}_{i+j} \hat{r}_j$

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	10775462	9159050	2581795	1380284	871723	589888	1275556	1643537	522288	645271	614036	336393	388070	1000053
1	11884473	10065073	3093379	2242467	1492272	995877	1163719	1993176	1247114	619533	718504	310931	608087	1102822
2	12758868	12110948	4081430	3597889	1540359	2849339	1056422	1550731	518345	823759	158023	299748	642179	1327062
3	13419185	10691706	4007080	2387385	1759759	991055	650689	1270744	606076	324085	282317	369345	635838	1315878
4	14484622	11786802	3959705	3109898	1465457	1140512	526079	1509184	597958	98371	546296	401205	691694	1433561
5	16217857	16020189	5515854	2399394	1732678	1733049	604095	621670	421763	648060	617720	454322	784412	1628091
6	19363970	16110879	6531044	1951440	2291509	743989	701399	1181296	880738	709035	676827	498520	861977	1791683
7	19613399	15933350	4781257	2106453	1447124	1933096	1542002	1884524	946647	763207	729602	538174	931895	1939825
8	21210290	18943798	7658792	3274449	1538707	1311305	1280249	2019959	1016161	820446	785463	580221	1006162	2097457
9	22725871	20516561	6237526	3472574	1267465	2102115	1537052	2428681	1223555	989335	948527	701695	1218573	2543930
10	21153719	23833794	9364579	4229022	2248657	2089046	1529727	2420633	1221277	988928	949515	703444	1223380	2557657
11	21740344	19845777	8890989	3831466	2265506	2107774	1545692	2449458	1237617	1003615	965014	715961	1246950	2610698
12	19870614	21076105	7167824	3640955	2156003	2008821	1475273	2341264	1184669	962071	926407	688312	1200526	2517125
13	21950030	21159535	7680870	3907259	2317072	2162037	1590104	2527166	1280589	1041474	1004316	747275	1305246	2740631

### 3.3 Bornhuetter-Ferguson

Il modello sottostante è:

$$C_{i,j} = \mu_i b_j \quad i \in \{0, \dots, t\} \quad j \in \{0, \dots, t\}$$

dove  $\mu_i = C_{i,t} = CP_i L_i$  e

$$\hat{b}_j = \left( \prod_{h=j}^{t-1} \hat{f}_h \right)^{-1} \quad j < t$$

dove gli  $\hat{f}_j$  sono gli stessi stimati con il *Chain Ladder*, e  $b_t = 1$ .

Il *loss ratio* di mercato è fisso per tutti gli anni di origine e pari a 0.85, e i premi di competenza  $PC_i$  sono dati. Nel modello

$$O\hat{L}L_i = \hat{\mu}_i(1 - \hat{b}_i)$$

e  $R = \sum_{i=1}^t O\hat{L}L_i$ . Si elencano di seguito le stime ottenute:

anno_orig	0	1	2	3	4	5	6	7	8	9	10	11	12	13
C_ultimo	33352956	36785642	42042132	41535989	44833804	50198632	59936701	60708750	65651558	70342688	65476456	67292217	61504900	67941252

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
bJJ	0.327	0.627	0.734	0.788	0.821	0.85	0.872	0.905	0.921	0.934	0.946	0.954	0.969	1
R_i	0.000	1157446.725	1926292.206	2252046.683	2956513.960	3964692.72	5704018.440	7782832.306	9826645.041	12619026.697	13849438.282	17896492.832	22946052.506	45699960

### 3.4 Fisher-Lange: metodo 2

Il modello sottostante è

- Per i pagamenti incrementali:

$$P_{i,j}^* = n_{ij}\sigma_j^* \quad i \in \{0, \dots, t\} \quad j \in \{0, \dots, t\}$$

- Per il numero di sinistri chiusi

$$\frac{n_{i,j}}{\sum_{h=j}^t n_{i,h}} = \psi_j$$

Per prima cosa si usano i tassi stimati con il metodo di Taylor per passare dai  $P_{i,j}$  ai  $P_{i,j}^*$ . Dopodichè si stimano:

$$\sigma_j^* = \frac{\sum_{i=0}^{t-j} P_{i,j}^*}{\sum_{i=0}^{t-j} n_{i,j}}$$

anno	0	1	2	3	4	5	6	7	8	9	10	11	12	13
sigma	2251	4292	9639	13790	17937	29639	34172	63414	33930	41453	36803	34182	17021	25001

Si stimano poi i  $\phi_{i,j}$ :

$$\hat{\psi}_j = \frac{\sum_{i=0}^{t-j} n_{i,j}}{\sum_{i=0}^{t-j} [n_i - \sum_{h=0}^{j-1} n_{i,h}]}$$

anno	0	1	2	3	4	5	6	7	8	9	10	11	12	13
phi	0.618	0.706	0.336	0.169	0.08	0.042	0.025	0.021	0.018	0.011	0.013	0.012	0.419	1

La proiezione dei numeri di sinistri si fa in modo iterativo:

$$\hat{n}_{i,j} = \hat{\psi} [n_i - \sum_{h=0}^{t-i} n_{i,h} - \sum_{h=t-i+1}^{j-1} \hat{n}_{i,h}]$$

Di seguito si mostra la matrice degli  $\hat{n}_{i,j}$  proiettati.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	14210	6896	630	231	119	116	54	31	29	20	19	9	32	40
1	14391	6718	674	252	193	102	37	39	24	20	15	12	30	46
2	13039	7490	992	390	184	87	40	36	28	12	17	10	996	1381
3	12795	6076	817	476	144	44	34	30	23	17	7	22	744	1033
4	12957	5320	992	252	91	35	32	24	24	8	30	27	930	1289
5	13863	6474	729	167	84	35	32	18	21	14	16	15	497	689
6	14921	5215	562	232	86	36	25	26	27	16	19	17	578	801
7	14577	4788	757	240	74	47	41	32	28	17	19	18	600	833
8	13928	5221	809	245	104	66	29	24	20	12	14	13	439	609
9	15115	5984	999	275	173	42	24	19	16	10	11	10	357	495
10	13218	6186	818	518	44	21	12	10	8	5	6	5	182	252
11	12200	4710	1579	174	69	33	19	15	13	8	9	8	285	396
12	10174	5553	388	129	51	25	14	11	10	6	7	6	213	295
13	10296	4354	611	203	80	39	22	18	15	9	11	10	334	464

Infine, si proiettano i  $\hat{P}_{i,j} = \hat{n}_{i,j} \hat{\sigma}_{i,j}$ , che sono espressi a valori correnti dell'anno t, e si adeguano rispetto all'inflazione usando i tassi stimati con il metodo di Taylor. Di seguito si mostrano i valori finali dei  $P_{i,j}$  espressi a valori correnti.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	30900604	24960290	7031519	3327979	1941226	1163257	2117341	2439700	763155	916974	812445	401553	447246	1000053
1	32387627	27412229	7458394	4993710	2942755	1653092	1727449	2912381	1772236	819719	857679	358344	608087	1261666
2	34748783	29200504	9088865	7095026	2556896	4229618	1543618	2203696	685834	983322	182120	299748	18594321	41620818
3	32354771	23809172	7901948	3962905	2612222	1448107	924674	1681351	723474	373504	282317	818190	15273587	34237753
4	32255551	23243535	6572856	4616397	2141291	1620747	696068	1801515	689139	98371	1207731	1122296	20981112	47100519
5	31981560	26592483	8187849	3505938	2462256	2293038	721109	716467	421763	637085	709485	660260	12361428	27790598
6	32142945	23915324	9543007	2773131	3031950	888101	808354	1181296	991999	813280	907026	845326	15849297	35683744
7	29114538	23281433	6794498	2787097	1727434	2227869	1542002	2238676	1133107	930322	1039074	969801	18209577	41057312
8	30991973	26920450	10133527	3908714	1773341	1311305	1101234	1797535	911153	749181	837976	783247	14728088	33255718
9	32295038	27145942	7445743	4002099	1267465	1359933	983723	1608067	816302	672169	752930	704778	13271775	30010787
10	27988980	28450433	10792563	4229022	859913	760391	550840	901759	458426	378031	424067	397522	7496634	16976268
11	25951479	22872015	8890989	2627781	1484357	1314482	953624	1563412	795946	657313	738428	693208	13091702	29689207
12	22900639	21076105	4107621	2151390	1217033	1079323	784162	1287459	656408	542865	610740	574168	10859210	24661894
13	21950030	20499273	7093499	3720687	2107848	1872064	1362091	2239567	1143493	947066	1067018	1004573	19026824	43273226

3.5 Confronto dei metodi

OLL per anno di origine: metodi a confronto

