Value-at-Risk: historical simulation

- approccio non parametrico: historical simulation
  - si usa la funzione di ripartizione empirica costruita da un campione casuale $l_1, \ldots, l_m$ osservato da L

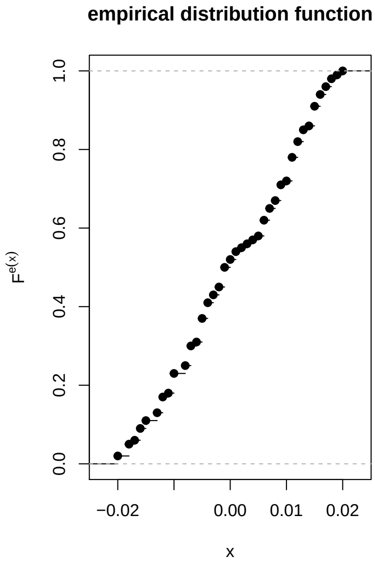
OSSERVATO OPPURE

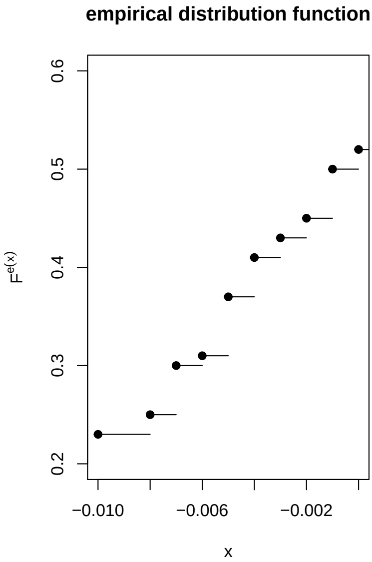
$$F_L^e(x) = \frac{\text{numero di } i : l_i \leq x}{m}$$

  - indicato con $l_{(1)} \leq l_{(2)} \leq \ldots \leq l_{(m)}$ la statistica d'ordine, se gli l_i sono tutti distinti riesce $F_L^e(l_{(i)}) = \frac{i}{m}$: salti in corrispondenza agli $l_{(i)}$ pari a $1/m$, costante tra due $\cancel{x_{(i)}}$ $l(i)$
  - se gli l_i non sono tutti distinti: se $l_{(i-1)} < l_{(i)} = l_{(i+1)} = \ldots = l_{(i+k-1)} < l_{(i+k)}$ il salto in corrispondenza a $l_{(i)}$ è pari a k/m, per il resto costante a tratti
  - metodo Monte Carlo: gli l_i sono simulati da un modello parametrico

SIMULATO

Value-at-Risk: historical simulation





VALUE-AT-RISK: HISTORICAL SIMULATION

FUNZIONALE DEFINITO SULL'INSIEME DELLE F. DI RIPARTIZIONE DI L

- Data una misura di rischio invariante per distribuzione $\rho(L) = \mathcal{G}_\rho(F_L)$,viene calcolata la sua versione empirica come

$$\rho^e(L) = \mathcal{G}_\rho(F_L^e)$$

- il Value-at-Risk è

$$\mathrm{VaR}_\alpha(L) = \inf\{x \in \mathbb{R} : \ F_L^e(x) \geq \alpha\}$$

  - ⋆ se αm è un intero, allora

$$\mathrm{VaR}_\alpha(L) = l_{(\alpha m)}$$

  - ⋆ in generale, ~~(necessariamente) un intero~~, ~~allora~~

$$\mathrm{VaR}_\alpha(L) = l_{(\lceil \alpha m \rceil)}$$

dove $\lceil x \rceil$ = minimo intero maggiore o uguale a x $\quad \lceil x \rceil - 1 < x \leq \lceil x \rceil$

HISTORICAL SIMULATION

es:

* $m = 1000$, $\alpha = 99.5\%$

$\alpha m = 995$ $\qquad VaR_{99.5\%}(L) = \ell_{(995)}$

$\ell_{(994)} \quad \ell_{(995)} \quad \ell_{(996)} \quad \ldots \quad \ell_{(1000)}$

$VaR_{99.5\%}$

* $m = 1000$, $\alpha = 99.55\%$

$\alpha m = 995.5$ $\qquad VaR_{99.55\%}(L) = \ell_{(996)}$

$\ell_{(994)} \quad \ell_{(995)} \quad \ell_{(996)} \quad \ldots \quad \ell_{(1000)}$

$VaR_{99.55\%}$

Value-at-Risk: historical simulation

SI VEDA LA FUNZIONE "QUANTILE" IN R

▷ Value-at-Risk: historical simulation

- ⋆ spesso si usa una variante che utilizza una media dei due valori $l_{(\lfloor \alpha m \rfloor)}$, $l_{(\lfloor \alpha m \rfloor + 1)}$ (se diversi), dove $\lfloor x \rfloor$ = massimo intero minore o uguale a x $\lfloor x \rfloor \le x < \lfloor x \rfloor + 1$
- ⋆ si pone

$$\mathrm{VaR}_\alpha(L) = (1-\gamma) l_{(\lfloor \alpha m \rfloor)} + \gamma l_{(\lfloor \alpha m \rfloor + 1)}$$

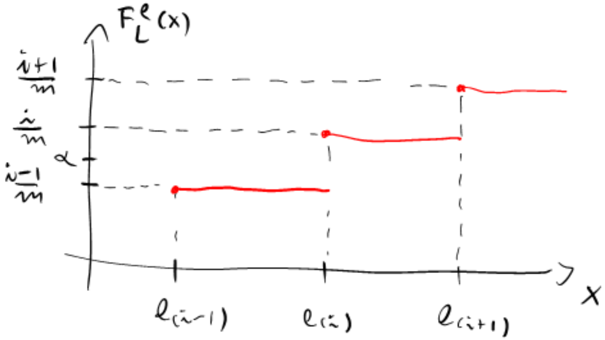
  dove $0 \le \gamma \le 1$ è un peso
- ⋆ classica scelta: $\gamma = \alpha m - \lfloor \alpha m \rfloor \rightsquigarrow \boxed{\text{quantile continuo in } \alpha}$
- ⋆ per $\gamma = 0$ se αm intero, $\gamma = 1$ se αm non intero, si ritorna alla definizione precedente

▷ Per l'expected shortfall, USANDO LA VERSIONE DI P. 312

$$\mathrm{ES}_\alpha(L) = \frac{1}{m(1-\alpha)} \left(\sum_{k=\lceil m\alpha \rceil + 1}^{m} l_{(k)} + (\lceil m\alpha \rceil - m\alpha) l_{(\lceil m\alpha \rceil)} \right),$$

dove il secondo termine additivo dentro parentesi viene a volte trascurato

INVERSA DI F^e



$F_L^{-1}(\alpha) = l_{(i)}$

$\iff$

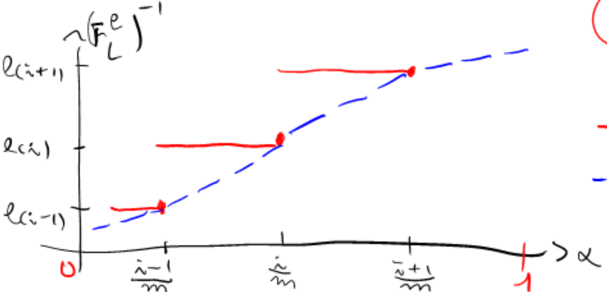
$\frac{i-1}{m} < \alpha \le \frac{i}{m}$

$\iff$

$i-1 < \alpha m \le i$

$\iff$

$i = \lceil \alpha m \rceil$



—— $(F^e_L)^{-1}$

- - - VERSIONE CONTINUA (INTERPOLAZIONE LINEARE)

Expected Shortfall: historical simulation

PERDITE IN % (EG OPPOSTO DEI RENDIMENTI)

- ESEMPIO: $l_1 = 1\%$, $l_2 = -2\%$, $l_3 = 0\%$, $l_4 = -1\%$, $l_5 = 2.5\%$, $l_6 = -1\%$, $l_7 = 3\%$, $l_8 = 0.5\%$, $l_9 = 1\%$, $l_{10} = 4\%$
  - ⋆ statistica d'ordine:

$$l_{(1)} = -2\%,\ l_{(2)} = l_{(3)} = -1\%,\ l_{(4)} = 0\%,\ l_{(5)} = 0.5\%,$$
$$l_{(6)} = l_{(7)} = 1\%,\ l_{(8)} = 2.5\%,\ l_{(9)} = 3\%,\ l_{(10)} = 4\%$$

  si trova $\mathrm{VaR}_{0.9}(L) = l_{(9)} = 3\%$ $\alpha = 0.9$ $\alpha n = 9$
  - ⋆ per $\beta > 0.9$, riesce $\mathrm{VaR}_\beta = 4\%$
  - ⋆ l'expected shortfall è

$$\mathrm{ES}_{0.9}(L) = \frac{1}{0.1}\int_{0.9}^{1} \overbrace{\mathrm{VaR}_\beta(L)}^{=4\% \text{ ECCETTO PER } \beta=0.9}\mathrm{d}\beta = 4\% = E[L|L > \mathrm{VaR}_{0.9}]$$

  - ⋆ si noti che $E[L|L \geq \mathrm{VaR}_{0.9}] = 3.5\%$ $= \dfrac{3\% \cdot 1 + 4\% \cdot 1}{2}$

ESEMPIO (CONTINUA)

$\alpha = 75\%$ $\quad m\alpha = 7.5$ $\quad \lceil m\alpha \rceil = 8$

$VaR_{75\%}(L) = \ell_{(8)} = 2.5\%$

$ES_{75\%}(L)$: VIA DEFINIZIONE

$$VaR_\beta(L) = \begin{cases} 2.5\% & 75\% \leq \beta \leq 80\% \\ 3\% & 80\% < \beta \leq 90\% \\ 4\% & \beta > 90\% \end{cases}$$

$$ES_{75\%}(L) = \frac{1}{1-0.75} \int_{0.75}^{1} VaR_\beta(L)\, d\beta = \frac{1}{0.25}\Big(2.5\% \cdot (0.8 - 0.75) + 3\% \cdot (0.9 - 0.8) + 4\% \, (1 - 0.9)\Big)$$

$$= 3.3\%$$

ESEMPIO (CONTINUA)

$ES_{75\%}(L)$: HISTORICAL SIMULATION

$$ES_{75\%}(L) = \frac{1}{m(1-\alpha)}\left(\sum_{u=\lceil m\alpha\rceil+1}^{m} l_{(u)} + (\lceil m\alpha\rceil - m\alpha)\cdot l_{\lceil m\alpha\rceil}\right)$$

$$= \frac{1}{10\cdot 0.25}\Big(\underset{3\%}{l_{(9)}} + \underset{4\%}{l_{(10)}} + (8-7.5)\cdot \underset{2.5\%}{l_{(8)}}\Big)$$

$$= 3.3\%$$

OSS: $E(L \mid L > VaR_{75\%}(L)) = 3.5\%$; $E(L \mid L \geq VaR_{75\%}(L)) = 3.17\%$

Misure di rischio

⊗ $\rho(L) = E(L) + \lambda\, VAR(L)$
PER λ GRANDE NON LA VERIFICA

- ▷ approccio assiomatico alle misure di rischio: stabilire proprietà ragionevoli che una misura di rischio $\rho : \mathcal{L} \to \mathbb{R}$ potrebbe soddisfare
  - ⋆ no rip-off:
    ⊗ $\rho(L) \leq$ estremo superiore di $L = \inf\{x \in \mathbb{R} : F_L(x) = 1\}$ per ogni $L \in \mathcal{L}$;
    "QUASI" TUTTE LE MISURE SODDISFANO QUESTA
    non si dovrebbe allocare più capitale della perdita massima OVVIO
  - ⋆ invarianza rispetto alla distribuzione: $\rho(L_1) = \rho(L_2)$ per ogni $L_1, L_2 \in \mathcal{L}$ tali che $F_{L_1} = F_{L_2}$;
    richiesta naturale, dato che osservazioni di L permettono di ricostruire solo F_L
  - ⋆ monotonia: $\rho(L_1) \leq \rho(L_2)$ per ogni $L_1, L_2 \in \mathcal{L}$ tali che $\text{Prob}(L_1 \leq L_2) = 1$;
    la perdita L_2 non è mai inferiore a $L_1 \Rightarrow$ richiede almeno tanto capitale quanto L_1

PIÙ BUSINESS => PIÙ RISCHIO => PIÙ CAPITALE

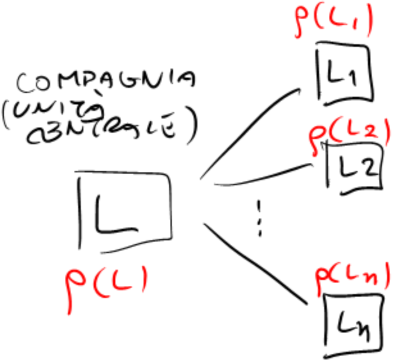
L_2 DOMINA STOCASTICAMENTE L_1 => $P(L_1 \leq L_2) = 1$

MISURE DI RISCHIO

- ▷ approccio assiomatico: stabilire proprietà ragionevoli che una misura di rischio potrebbe soddisfare
  - ⋆ invarianza per traslazioni: $\rho(L+c) = \rho(L) + c$ per ogni $L \in \mathcal{L}$, $c \in \mathbb{R}$; IL $\rho(L)$ È CAPITALE NECESSARIO!
    aggiungendo una perdita certa c ad L, si deve aumentare il capitale dello stesso ammontare $\Rightarrow \rho(L - \rho(L)) = 0$
  - ⋆ sub-additività: $\rho(L_1 + L_2) \leq \rho(L_1) + \rho(L_2)$ per ogni $L_1, L_2 \in \mathcal{L}$
    beneficio della diversificazione; beneficio di una fusione; sistema di allocazione di capitale può essere decentralizzato VEDI DOPO
  - ⋆ positiva omogeneità: $\rho(\lambda L) = \lambda \rho(L)$ per ogni $\lambda \geq 0$, $L \in \mathcal{L}$;
    non si penalizza per rischio di concentrazione e liquidità
  - ⋆ convessità: $\rho(\lambda L_1 + (1-\lambda) L_2) \leq \lambda \rho(L_1) + (1-\lambda)\rho(L_2)$ per ogni $0 < \lambda < 1$ e $L_1, L_2 \in \mathcal{L}$;
    diversificazione su composizioni di portafogli

SUBADDITIVITÀ

("DECENTRALIZED ALLOCATION OF CAPITAL")



LINEA DI BUSINESS/RAMO/TRADER/UNITÀ 1

" " 2

" " n

$$\rho(L) \le \underbrace{\rho(L_1)+\rho(L_2)+\dots+\rho(L_n)}_{\text{SE CALCOLO IL CAPITALE COME SOMMA} \Rightarrow \text{CONSERVATIVO}}$$

CONVESSITÀ

FONDO	PERDITA	QUOTA	
1	L_1	λ_1	5%
2	L_2	λ_2	15%
$\vdots$			$\vdots$
n	L_n	λ_n	10%

$\sum \lambda_i = 1$

$$\rho(\lambda_1 L_1 + \dots + \lambda_n L_n) \leq$$

$$\leq \lambda_1 \rho(L_1) + \dots + \lambda_n \rho(L_n)$$

$$= \underbrace{\rho(\lambda_1 L_1) + \dots + \rho(\lambda_n L_n)}_{\text{SOMMA DEI CAPITALI SEPARATI}} \quad \text{SE POS. OMOGENEA}$$

$\rightarrow$ CAPITALE DEL FONDO COSÌ COSTRUITO

SUBADDITIVITÀ

FUSIONE (MERGER) DI 2 (O PIÙ) COMPAGNIE / UNITÀ / FONDI

$$\underbrace{\rho(L_1+L_2)}_{\text{CAPITALE POST-FUSIONE}} \le \underbrace{\rho(L_1)+\rho(L_2)}_{\text{SOMMA DEI CAPITALI PRE-ALLOCAZIONE}}$$

FUSIONE => DIVERSIFICAZIONE

=> RIDUZIONE DI CAPITALE

POSITIVA OMOGENEITÀ

$$\rho(\lambda L) = \lambda \rho(L) \quad \text{PER OGNI } \lambda \geq 0, \; L \in \mathcal{L}$$

ESEMPIO: L = FONDO D'INVESTIMENTO

λL = STESSO FONDO MA CON ESPOSIZIONI AUMENTATE $\quad \lambda > 1$

		L	$10 \cdot L$
STOCK	1	20M€	200M€
	2	30M€	300M€
	3	50M€	500M€

$$\underbrace{\rho(10 \cdot L)}_{\text{RIFLETTE}} = 10 \cdot \rho(L)$$

RIFLETTE

* CONCENTRAZIONE?
* LIQUIDITÀ?

Misure di rischio

- una misura di rischio che soddisfa la monotonia, invarianza per traslazioni, positiva omogeneità e sub-additiva si dice coerente
- una misura di rischio che soddisfa la monotonia, invarianza per traslazioni e convessità si dice convessa → DIVERSIFICAZIONE SENZA POSITIVA OMOGENEITÀ
- ogni misura di rischio positivamente omogenea è subadditiva se e solo se verifica la convessità ⇝ ogni misura di rischio coerente è convessa
- Il Value-at-Risk soddisfa tutte le proprietà (in particolare monotonia, invarianza per traslazioni e positiva omogeneità) eccetto la subadditività e quindi neanche la convessità ⇝ il Value-at-Risk non è coerente
  - ⋆ il Value-at-Risk diventa coerente se ci si restringe a certi insiemi di variabili aleatorie (normali o, più in generale, ellittiche)
- L'expected shortfall verifica tutte le proprietà viste prima È COERENTE È CONVESSA
- le misure di rischio basate su scenari generalizzati ~~(con pesi tutti unitari)~~ sono coerenti

MISURE DI SCENARIO GENERALIZZATE

$$\rho(L) = \sup \{ E^Q[L] \, , \, Q \in \mathcal{P} \}$$

INSIEME DI PROB. (SCENARI)

2) INVARIANZA PER TRASLAZIONE

$$\rho(L+c) = \sup \{ E^Q(L+c) \, , \, Q \in \mathcal{P} \}$$

$$= \sup \{ E^Q(L) + c \, , \, Q \in \mathcal{P} \}$$

$$= \sup \{ E^Q(L) \, , \, Q \in \mathcal{P} \} + c$$

$$= \rho(L) + c$$

-- CONTINUA

3) POSITIVA OMOGENEITÀ

$$\rho(\lambda L) = \sup\{E^Q(\lambda L),\ Q \in \mathcal{P}\}$$
$$= \sup\{\lambda E^Q(L),\ Q \in \mathcal{P}\}$$
$$\overset{\lambda \geq 0}{=} \lambda \cdot \sup\{E^Q(L),\ Q \in \mathcal{P}\}$$
$$= \lambda \cdot \rho(L)$$

... CONTINUA

4) SUBADDITIVITÀ

$$\rho(L_1+L_2) = \sup\{E^Q(L_1+L_2),\ Q\in\mathcal{P}\}$$
$$= \sup\{E^Q(L_1)+E^Q(L_2),\ Q\in\mathcal{P}\}$$
$$\leq \sup\{E^Q(L_1),\ Q\in\mathcal{P}\}$$
$$+ \sup\{E^Q(L_2),\ Q\in\mathcal{P}\}$$
$$= \rho(L_1)+\rho(L_2)$$

(ESEMPIO: $\max(a_1+b_1, a_2+b_2, \dots, a_n+b_n) \leq$
$\leq \max\{a_1, \dots, a_n\} + \max\{b_1, \dots, b_n\}$)

... CONTINUA

1) MONOTONIA

RICHIEDE CHE Q ASSOLUTAMENTE CONTINUA RISPETTO A P $(\Omega, \mathcal{F}, P)$ PER OGNI $Q \in \mathcal{P}$

($P(A)=0 \Rightarrow Q(A)=0$) PER OGNI $A \in \mathcal{F}$

$$P(L_1 \leq L_2) = 1 \Rightarrow Q(L_1 \leq L_2) = 1 \quad \forall Q \in \mathcal{P}$$

$$\Rightarrow E^Q(L_1) \leq E^Q(L_2) \quad \forall Q$$

$$\Rightarrow \rho(L_1) = \sup\{E^Q(L_1), Q \in \mathcal{P}\} \leq$$

$$\leq \sup\{E^Q(L_2), Q \in \mathcal{P}\} = \rho(L_2)$$

SE ρ È POS. OMOGENEA

È SUBADDITIVA $\Leftrightarrow$ È CONVESSA

$\Rightarrow$

$$\rho(\lambda L_1 + (1-\lambda)L_2) \leq \rho(\lambda L_1) + \rho((1-\lambda)L_2)$$
$$= \lambda\rho(L_1) + (1-\lambda)\rho(L_2)$$

$\Leftarrow$

$$\rho(L_1 + L_2) = \rho\left(\frac{1}{2}(2L_1) + \frac{1}{2}(2L_2)\right)$$
$$\leq \frac{1}{2}\rho(2L_1) + \frac{1}{2}\rho(2L_2)$$
$$= \cancel{\frac{1}{2}}\cancel{2}\rho(L_1) + \cancel{\frac{1}{2}}\cancel{2}\rho(L_2)$$

VALUE AT RISK

1) **MONOTONIA**: $P(L_1 \leq l_2) = 1$

$\Rightarrow$ L_1 DOMINATA STOCASTICAMENTE DA l_2

$\Rightarrow$ $VaR_\alpha(L_1) \leq VaR_\alpha(L_2)$

2) **INVARIANZA PER TRASLAZIONE**

$$VaR_\alpha(\underbrace{L + c}_{f(L)}) = VaR_\alpha(L) + c$$

$f \uparrow$

3) **POSITIVA OMOGENEITÀ**

$$VaR_\alpha(\underbrace{\lambda L}_{f(L)}) = \lambda \cdot VaR_\alpha(L)$$

$\lambda \geq 0$ $\quad$ $f \uparrow$

EXPECTED SHORTFALL

1) MONOTONIA: OK

2) INVARIANZA PER TRASLAZIONE

$$ES_\alpha(\underbrace{L+c}) = ES_\alpha(L)+c$$

TRASFORMAZIONE
LINEARE CRESCENTE

3) POSITIVA OMOGENEITÀ

$$ES_\alpha(\underbrace{\lambda L}) = \lambda ES_\alpha(L)$$

$\lambda \geq 0$ TRASFORMAZIONE LINEARE CRESCENTE

Subadditività dell'Expected Shortfall

- Prova della subadditività nel caso in cui L_1, L_2 e $L = L_1 + L_2$ siano continue ($\mathrm{ES}_\alpha(L_i) = \frac{E[L_i 1_{L_i \geq \mathrm{VaR}_\alpha(L_i)}]}{1-\alpha}$, $i = 1,\ 2$)
  - $I_i = 1_{\{L_i \geq \mathrm{VaR}_\alpha(L_i)\}}$, $i = 1,\ 2$, $I_{12} = 1_{\{L \geq \mathrm{VaR}_\alpha(L)\}}$
  - si osservi che

$$(1-\alpha)[\mathrm{ES}_\alpha(L_1)+\mathrm{ES}_\alpha(L_2)-\mathrm{ES}_\alpha(L)] = E[L_1(I_1-I_{12})]+E[L_2(I_2-I_{12}]$$

  - riesce

$$L_i(I_i - I_{12}) \geq \mathrm{VaR}_\alpha(L_i)(I_1 - I_{12}),\ i = 1, 2,$$

    (considerare separatamente i casi $L_i \geq \mathrm{VaR}_\alpha(L_i)$ e $L_i \geq \mathrm{VaR}_\alpha(L_i)$)
  - concludere che

$$\mathrm{ES}_\alpha(L_1) + \mathrm{ES}_\alpha(L_2) \leq \mathrm{ES}_\alpha(L)$$

PROVA DELLA SUBADDITIVITÀ (CASO CONTINUO) DI ES_α

IPOTESI: $L_1, L_2, L_1 + L_2$ CONTINUE

TESI: $ES_\alpha(L_1+L_2) \leq ES_\alpha(L_1) + ES_\alpha(L_2)$

(OSSERVAZIONE: IL RISULTATO VALE ANCHE SENZA L'IPOTESI)

DIMOSTRAZIONE:

$$I_i = 1_{\{L_i \geq VaR_\alpha(L_i)\}} \qquad i=1,2$$

$$I_{12} = 1_{\{L_1 + L_2 \geq VaR_\alpha(L_1+L_2)\}}$$

... CONTINUA

$$ES_\alpha(L_i) = E[L_i \mid L_i \geq VaR_\alpha(L_i)]$$

$$= \frac{E[L_i I_i]}{1-\alpha}$$

$$ES_\alpha(L_1+L_2) = \frac{E[(L_1+L_2) I_{12}]}{1-\alpha}$$

GRAZIE ALLA CONTINUITÀ

... CONTINUA

$$(1-\alpha)\left\{ES_\alpha(L_1)+ES_\alpha(L_2)-ES_\alpha(L_1+L_2)\right\}$$

$$=E\left[L_1 I_1 + L_2 I_2 - (L_1+L_2) I_{12}\right] \quad (*)$$

$$=E\left[L_1(I_1-I_{12})\right]+E\left[L_2(I_2-I_{12})\right]$$

DUE POSSIBILITÀ:

1) $L_1 \geq VaR_\alpha(L_1)$, CIOÈ $I_1=1$

$$\Rightarrow L_1(\underbrace{\overset{=1}{I_1}-I_{12}}_{\geq 0}) \geq VaR_\alpha(L_1)\cdot(I_1-I_{12})$$

... CONTINUA

2) $L_1 < VaR_\alpha(L_1)$, CIOÈ $I_1 = 0$

$$L_1 \underbrace{(\overset{=0}{I_1} - I_{12})}_{\leq 0} \geq VaR_\alpha(L_1)(I_1 - I_{12})$$

$\to$ IN OGNI CASO, $L_1(I_1 - I_{12})$
$\geq VaR_\alpha(L_1)(I_1 - I_2)$

$\to$ STESSO PER L_2!

... CONTINUA

RIPRENDENDO DA (*),

$$(1-\alpha)\{ES_\alpha(L_1)+ES_\alpha(L_2)-ES_\alpha(L_1+L_2)\}$$

$$= E[L_1(I_1-I_{12})] + E[L_2(I_2-I_{12})]$$

$$\geq VaR_\alpha(L_1)\cdot \underbrace{E[I_1-I_{12}]}_{=\,E[I_1]-E[I_{12}]} + VaR_\alpha(L_2)\cdot \underbrace{E[I_2-I_{12}]}_{=0} = 0$$

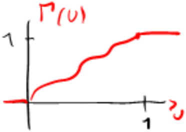
$$E[I_1] = \underbrace{P(L_1 \geq VaR_\alpha(L_1))}_{=1-\alpha}, \qquad E[I_{12}] = \underbrace{P(L_1+L_2 \geq VaR_\alpha(L_1+L_2))}_{=1-\alpha}$$

$$=0$$

$$\Longrightarrow ES_\alpha(L_1)+ES_\alpha(L_2) \geq ES_\alpha(L_1+L_2)$$

Misure di rischio distorte

- Il Value-at-Risk e l'expected shortfall possono essere generalizzati al modo seguente, introducendo le misure di rischio distorte: per una perdita L e data una funzione di ripartizione $\Gamma : [0,1] \to [0,1]$ (Γ non decrescente, continua a dx, $\Gamma(0) = 0$, $\Gamma(1) = 1$), si pone



Γ UNIFORME

$\Rightarrow \rho_\Gamma(L) = E[L]$

$$\rho_\Gamma(L) = \int_0^1 \mathrm{VaR}_\beta(L) \mathrm{d}\Gamma(\beta)$$

  - ⋆ idea: pesare i capitali a ogni livello di confidenza con la misura Γ
  - ⋆ VaR_α: Γ concentrata in α

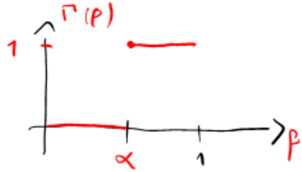
- quando Γ è assolutamente continua con densità γ,

$$\rho_\gamma(L) = \int_0^1 \mathrm{VaR}_\beta(L)\gamma(\beta)\mathrm{d}\beta$$

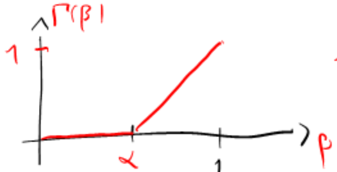
  - ⋆ Expected shortfall a livello α: $\gamma(\beta) = \frac{1}{1-\alpha} 1_{[\alpha,1]}(\beta)$
  - ⋆ $\gamma(\beta) = \frac{\lambda \exp(\lambda\beta)}{\exp(\lambda)-1}$ per $\lambda > 0$
  - ⋆ se γ è non decrescente (misure di rischio spettrali), si può provare che ρ_γ è coerente

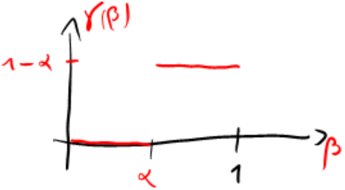
MISURE DI RISCHIO DISTORTE

VaR_α



ES_α



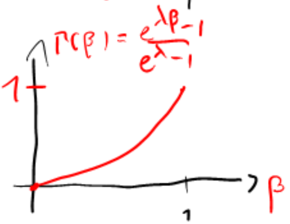


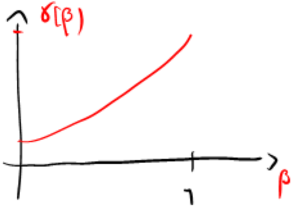
ESPONENZIALE

$$\gamma(\beta) = \frac{\lambda e^{\lambda\beta}}{e^{\lambda}-1}$$

$\downarrow$

AUMENTA IL PESO SULLA CODA (ESISTE?)





γ NON DECRESCENTE $\Rightarrow$ COERENTE

Misure di rischio basate su perdite

EVITA DI ASSUMERE
LA POSITIVA OMOGENEITÀ!
* LIQUIDITÀ, CONCENTRAZIONE

- ▷ Una misura di rischio convessa non (positivamente omogenea) ⇝ non coerente
- ▷ $\ell : \mathbb{R} \to \mathbb{R}$ crescente convessa (funzione di perdita) e $k \in \mathbb{R}$ una soglia di perdita; sia $E[|\ell(L)|] < +\infty$ per ogni $L \in \mathcal{L}$; definiamo

$$\rho(L) \equiv \rho_{\ell,k}(L) = \inf\{x \in \mathbb{R} |\, E[\ell(L - x)] \le \ell(k)\}$$

- ▷ ρ è convessa ma non necessariamente coerente
  - ⋆ $\ell(z) = \exp(\lambda z)$, $\lambda > 0$
  - ⋆ $\rho(L) = \frac{1}{\lambda} \log E[\exp(\lambda L)] - k$
  - ⋆ per $k = 0$ e ~~$\lambda > 1$ riesce $\rho(\lambda L) > \lambda\rho(L)$~~ (si usi la disuguaglianza di Jensen applicata a $z^{\cancel{\lambda}}$)

z^{μ}

$\mu > 1$, RIESCE
$\underline{\rho(\mu L) > \mu\, \rho(L)}$

MISURA DI RISCHIO CONVESSA NON COERENTE

$$\rho(L) = \inf \{ x \in \mathbb{R} \mid E[\ell(L-x)] \leq \ell(K) \}$$

$$\mathcal{L} = \{ L \text{ on } (\Omega, \mathcal{F}, P) \mid : E[|\ell(L)|] < +\infty \}$$

IDEA : ℓ = FUNZIONE DI PERDITA ("OPPOSTO" DI UNA FUNZIONE DI UTILITÀ $u(z) := -\ell(-z)$ È CRESCENTE CONCAVA)

$\ell(K)$ = MASSIMO LIVELLO DI PERDITA ACCETTABILE

$\rho(L)$ = MINIMO CAPITALE ALLOCATO TALE CHE LA PERDITA NETTA È ACCETTABILE!

... CONTINUA

MONOTONIA

$$P(L_1 \leq L_2) = 1$$

$$\Rightarrow E[\ell(L_1 - x)] \leq E[\ell(L_2 - x)] \quad \forall x$$

$$\Rightarrow \rho(L_1) \leq \rho(L_2)$$

INVARIANZA PER TRASLAZIONI

$y = x - c$

$$\rho(L + c) = \inf\{x \in \mathbb{R} \mid E[\ell(L + c - x)] \leq \ell(K)\}$$

$$= \inf\{y + c \in \mathbb{R} \mid E[\ell(L - y)] \leq \ell(K)\}$$

$$= \rho(L) + c$$

... CONTINUA

CONVESSITÀ : L_1, L_2, $0 \leq \lambda \leq 1$

$$\tilde{L}_i = L_i - \rho(L_i)$$

$$\rho(\tilde{L}_i) = \rho(L_i) - \rho(L_i) = 0$$

(INVARIANZA PER TRASLAZIONI)

MOSTRIAMO CHE

$$\underbrace{\rho(\lambda \tilde{L}_1 + (1-\lambda)\tilde{L}_2)}_{} \leq 0$$

$$\inf \{ x \mid E[\ell(\lambda \tilde{L}_1 + (1-\lambda)\tilde{L}_2 - x)] \leq \ell(x) \}$$

$$\forall x > \underset{\inf\{--\}}{\overset{}{0}} , \quad E[\ell(\tilde{L}_i - x)] \leq \ell(x)$$

... CONTINUA. SEGUE CHE

$$E\Big[\ell\big(\underbrace{\lambda\tilde{L}_1+(1-\lambda)\tilde{L}_2-x}_{=\lambda(\tilde{L}_1-x)+(1-\lambda)(\tilde{L}_2-x)}\big)\Big]$$

CONVESSITÀ DI ℓ

$$\leq \lambda\underbrace{E\big[\ell(\tilde{L}_1-x)\big]}_{\leq \ell(u)}+(1-\lambda)\underbrace{E\big[\ell(\tilde{L}_2-x)\big]}_{\leq \ell(u)} \leq \ell(u)$$

$$\Rightarrow \rho(\lambda\tilde{L}_1+(1-\lambda)\tilde{L}_2)\leq 0$$

PER FINIRE:

$$\rho(\lambda L_1+(1-\lambda)L_2)-\lambda\rho(L_1)-(1-\lambda)\rho(L_2)$$

$$= \rho(\lambda\tilde{L}_1+(1-\lambda)\tilde{L}_2)\leq 0$$

INVARIANZA PER TRASLAZIONI

ESEMPIO DI ρ NON POSITIVAMENTE OMOGENEA

$$\ell(z) = e^{\lambda z}, \quad \lambda > 1$$

$$\Rightarrow \boxed{\rho(L) = \frac{1}{\lambda} \log E\left[e^{\lambda L}\right] - K}$$

$$\rho(L) = \inf\Big\{ x \;\Big|\; \underbrace{E\left[e^{\lambda(L-x)}\right] \leq e^{\lambda K}} \Big\}$$

$$E\left[e^{\lambda L}\right] e^{-\lambda K} \leq e^{\lambda x}$$

$$\frac{1}{\lambda} \log E\left[e^{\lambda L}\right] - K \leq x$$

... CONTINUA

$\mu > 1$, $K = 0$

$$\rho(\mu L) = \frac{1}{\lambda} \log E\left[e^{\lambda(\mu L)}\right] =$$

$$= \frac{1}{\lambda} \log E\left[\underbrace{\left(e^{\lambda L}\right)^{\mu}}_{z^{\mu}}\right] >$$

z^{μ} CONVESSA SE $\mu > 1$

$$> \frac{1}{\lambda} \log E\left[e^{\lambda L}\right]^{\mu}$$

$$= \frac{\mu}{\lambda} \log E\left[e^{\lambda L}\right] = \mu \cdot \rho(L)$$

Value-at-Risk e subadditività

- la proprietà di subadditivita vale per il Value-at-Risk se $(L_1,\ L_2)$ è normale bivariata e $\alpha > 50\%$
- un vettore $\mathbf{X} = [X_1, \ldots, X_d]^\top$ ha distribuzione normale multivariata se

$$a_1 X_1 + \ldots + a_d X_d \sim N(\mu, \sigma^2)$$

  per qualche $\mu \in \mathbb{R}$, $\sigma^2 \geq 0$, per ogni $a_1, \ldots, a_d \in \mathbb{R}$ (una costante ha distribuzione $N(\mu, 0)$)
- se la matrice di varianza-covarianza $\Sigma = E[(\mathbf{X} - \mu)(\mathbf{X} - \mu)^\top]$ è invertibile ($\mu = E(\mathbf{X})$), allora $\mathbf{X}$ ha densità congiunta

$$f(\mathbf{x}) = (2\pi|\Sigma|)^{-d/2} \exp\left(-\frac{1}{2}(\mathbf{x} - \mu)^\top \Sigma^{-1} (\mathbf{x} - \mu)\right)$$

VAR È ADDITIVO NEL CASO

NORMALE (SE $\alpha \geq 50\%$)

$$(L_1, L_2) \sim N^{(2)}\left(\begin{bmatrix}\mu_1 \\ \mu_2\end{bmatrix}, \begin{bmatrix}\sigma_1 & \rho\sigma_1\sigma_2 \\ & \sigma_2\end{bmatrix}\right)$$

$$\Rightarrow VaR_\alpha(L_i) = \mu_i + \sigma_i \Phi^{-1}(\alpha) \qquad i = 1, 2$$

$$L_1 + L_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2)$$

$$\Rightarrow VaR_\alpha(L_1 + L_2) = \mu_1 + \mu_2 + \sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2} \cdot \Phi^{-1}(\alpha)$$

-- VaR NEL CASO NORMALE

$$\underbrace{VaR_\alpha(L_1+L_2)}_{\cancel{\mu_1}+\cancel{\mu_2}+\sqrt{\sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2}\,\Phi^{-1}(\alpha)} \leq \underbrace{VaR_\alpha(L_1)+VaR_\alpha(L_2)}_{\cancel{\mu_1}+\sigma_1\Phi^{-1}(\alpha)+\cancel{\mu_2}+\sigma_2\Phi^{-1}(\alpha)}$$

$\Phi^{-1}(\alpha) \geq 0 \quad (\alpha \geq 0.5)$

$$\iff \sqrt{\sigma_1^2+\sigma_2^2+2\rho\sigma_1\sigma_2} \leq \sigma_1+\sigma_2$$

$$\iff \cancel{\sigma_1^2}+\cancel{\sigma_2^2}+\cancel{2}\rho\,\cancel{\sigma_1\sigma_2} \leq \cancel{\sigma_1^2}+\cancel{\sigma_2^2}+\cancel{2\sigma_1\sigma_2}$$

$\iff \rho \leq 1$ SEMPRE VERA.

NORMALE MULTIVARIATA

COSTRUZIONE: $Z = \begin{bmatrix} Z_1 \\ \vdots \\ Z_d \end{bmatrix}$ IID $N(0,1)$

μ VETTORE DELLE MEDIE

Σ MATRICE VARIANZE - COVARIANZE (SIMMETRICA, SEMI DEFINITA POSITIVA)

ESISTE C MATRICE TRIANGOLARE (DI CHOLESKI) : $\Sigma = C \cdot C^T$

$\boxed{X = \mu + C \cdot Z}$ HA DISTRIBUZIONE MULTIVARIATA NORMALE MEDIA μ. COVARIANZA Σ

... NORMALE MULTIVARIATA

$$\mathrm{cov}(X) = \mathrm{cov}(\mu + C \cdot Z) = C^T \mathrm{cov}(Z)\, C$$

$$= C^T I_d\, C = C^T C = \Sigma$$

$$X_1 = \mu_1 + c_{11} Z_1$$

$$X_2 = \mu_1 + c_{21} Z_1 + c_{22} Z_2$$

$$\vdots$$

$$X_i = \mu_i + c_{i1} Z_1 + c_{i2} Z_2 + \cdots + c_{ii} Z_i$$

$$\vdots$$

$$X_d = \mu_d + c_{d1} Z_1 + c_{d2} Z_2 + \cdots + c_{dd} Z_d$$

Value-at-Risk e subadditività

- La subadditività del Value-at-Risk si estende ad altre famiglie di distribuzioni multivariate, in particolare alle distribuzioni ellittiche che hanno in comune diverse proprietà con la normale multivariata:
  1) ⋆ generalizzano al caso multivariato le proprietá di simmetria della normale
  2) ⋆ distribuzioni marginali di distribuzioni ellittiche sono ancora ellittiche
  3) ⋆ trasformazioni affini di distribuzioni ellittiche sono ancora ellittiche
  4) ⋆ distribuzioni condizionate di distribuzioni ellittiche sono ancora ellittiche

Value-at-Risk e subadditività

▷ un vettore $\mathbf{X} = [X_1, \ldots, X_d]^\top$ ha distribuzione sferica se

$$UX \sim X$$

per ogni matrice ortogonale U (cioè tale che $UU^\top = U^\top U = I_d$)

- ⋆ se $\mathbf{X}$ ha densità congiunta f, allora è sferica se e solo se

$$f(\mathbf{x}) = \widetilde{f}(\underbrace{x_1^2 + \ldots + x_d^2}_{x^\top x}), \quad \text{per ogni } \mathbf{x} \in \mathbb{R}^d$$

  dove $\widetilde{f} : [0, +\infty) \to [0, \infty)$, cioè la densità è costante sulle sfere. DI CENTRO L'ORIGINE

- ⋆ t_ν multivariata, con densità

$$f(\mathbf{x}) = C_\nu \left(1 + \frac{\mathbf{x}^\top \mathbf{x}}{\nu}\right)^{-(d+\nu)/2}$$

- ⋆ logistica multivariata, con densità

$$f(\mathbf{x}) = C_{\cancel{\nu}} \frac{\exp(-\mathbf{x}^\top \mathbf{x})}{[1 + \exp(-\mathbf{x}^\top \mathbf{x})]^2}$$

DISTRIBUZIONI SFERICHE

U ORTOGONALE; ESEMPI

* $U = \begin{bmatrix} \pm 1 & 0 & & 0 \\ 0 & \pm 1 & & \\ \vdots & & & \\ 0 & & 0 & \pm 1 \end{bmatrix}$ $\quad UX = \begin{bmatrix} \pm X_1 \\ \pm X_2 \\ \vdots \\ \pm X_d \end{bmatrix} \sim \begin{bmatrix} X_1 \\ \vdots \\ X_d \end{bmatrix}$

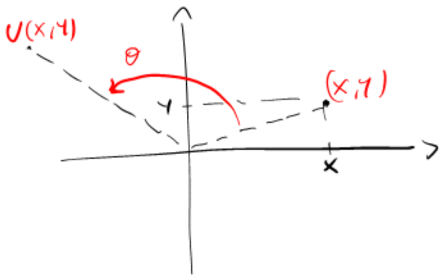
$\Rightarrow X_i$ SONO SIMMETRICHE

* $d=2$, ROTAZIONE DI UN ANGOLO PARI A θ

$$U = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

θ	TRASFORMAZIONE
0	$(x,y) \to (x,y)$
$\pi/2$	$(-y,x)$
π	$(-x,-y)$
$3\pi/2$	$(x,-y)$

... DISTRIBUZIONI SFERICHE

* 

* $d=1$: X DISTRIBUZIONE SFERICA $\iff$ X SIMMETRICA $(X \sim -X)$

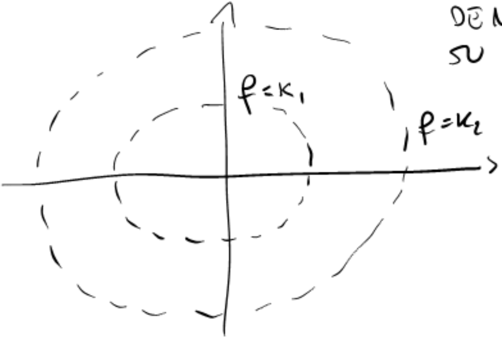
* $d>1$: ESTENDE L'IDEA DI SIMMETRIA

... DISTRIBUZIONI SFERICHE

DENSITÀ CONGIUNTA

$$f(x_1 \text{———} x_d) = \tilde{f}(x^T x) = \tilde{f}(\|x\|^2)$$

$$= \tilde{f}(x_1^2 + \text{———} + x_d^2)$$



DENSITÀ COSTANTE SU CERCHI CON CENTRO $\underline{0}$

... DISTRIBUZIONI SFERICHE

$\underline{X} = [X_1 \_\_ X_d]^T$ con distribuzione sferica

* (X_i) ugualmente distribuite

$$U = \begin{array}{c|c} & \begin{array}{ccccccc} & & i & & & j & \end{array} \\ \begin{array}{c} \\ i \\ \\ j \\ \end{array} & \begin{bmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & 0 & & & 1 & \\ & & & 1 & & & \\ & & & & 1 & & \\ & & & & & 1 & \\ & & 1 & & & 0 & \\ & & & & & & 1 \end{bmatrix} \end{array}$$

$$UX = \begin{array}{c} \\ i \\ \\ j \\ \end{array} \begin{bmatrix} X_1 \\ \vdots \\ X_j \\ \vdots \\ X_i \\ \vdots \end{bmatrix} \sim \begin{bmatrix} X_i \\ \\ X_j \end{bmatrix}$$

$$\Rightarrow X_i \sim X_j$$

* (X_i) non correlate

$$U = \begin{array}{c|c} & \begin{array}{cccc} & & i & \end{array} \\ i & \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & -1 & \\ & & & 1 \\ & & & & \ddots \\ & & & & & 1 \end{bmatrix} \end{array} \qquad UX = \begin{bmatrix} X_1 \\ -X_i \\ \vdots \\ X_d \end{bmatrix} \sim \begin{bmatrix} X_1 \\ \vdots \\ X_i \\ \vdots \\ X_d \end{bmatrix}$$

$$\Rightarrow \operatorname{cov}(X_i, X_j) = \operatorname{cov}(-X_i, X_j) = -\operatorname{cov}(X_i, X_j) = 0$$

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_d \end{bmatrix}$$

NORMALE MULTIVARIATA

$\mu = E[X] = 0$ $\quad cov(X) = \sigma^2 I_d$

$\Rightarrow$ X SFERICA

$UX \sim$ NORMALE MULTIVARIATA

U ORTOGONALE

$E[UX] = 0$

$cov(UX) = U\, cov(X)\, U^T$

$= \sigma^2 U I_d U^T = \sigma^2 I_d$

Value-at-Risk e subadditività

ES: NORMALE MULTIVARIATA

- un vettore $\mathbf{X} = [X_1, \ldots, X_d]^\top$ ha distribuzione ellittica se

$$\mathbf{X} \sim \mu + AY$$

CON Y CON DISTRIBUZIONE SFERICA

"CAMBIO DI LOCAZIONE μ E SCALA A" O "TRASFORMAZIONE AFFINE"

dove $\mu \in \mathbb{R}^d$ e A è una matrice $d \times d$

- se $L_1 = a^\top \mathbf{X}$, $L_2 = b^\top \mathbf{X}$ $(a, b, \in \mathbb{R}^d)$
  - ⋆ subadditività del VaR : per $\alpha \geq 0.5$

$$\mathrm{VaR}_\alpha(L_1 + L_2) \leq \mathrm{VaR}_\alpha(L_1) + \mathrm{VaR}_\alpha(L_2)$$

  - ⋆ Value-at-Risk è consistente con la varianza

$$\mathrm{VaR}_\alpha(L_1) - E[L_1] \leq \mathrm{VaR}_\alpha(L_2) - E[L_2] \Leftrightarrow var(L_1) \leq var(L_2)$$

⇝ ottimizzazione di portafoglio media-varianza o media-VaR coincidono

DISTRIBUZIONI ELLITTICHE

PROPRIETÀ

$\underline{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_d \end{bmatrix}$ CON DISTRIBUZIONE ELLITTICA

1) $\underline{X}' = \begin{bmatrix} X_{i_1} \\ \vdots \\ X_{i_m} \end{bmatrix}$ (MARGINALE DI $\underline{X}$) HA DISTRIBUZIONE ELLITTICA

$m < d$

2) $b + BX$ (b VETTORE, B MATRICE) HA DISTRIBUZIONE ELLITTICA

3) $X = \begin{bmatrix} X_I \\ X_J \end{bmatrix}$ $(X_I \mid X_J = z)$ HA DISTRIBUZIONE ELLITTICA

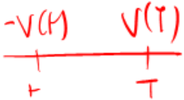
PROFITTO/PERDITA

- Sia $V(u)$ il valore/prezzo di un'attività al tempo u
  - stocks
  - bonds
  - commodities (merci)
  - derivati
  - ...
  - portafoglio di attività
- consideriamo il periodo $[t, T]$ (holding period) di lunghezza $T - t$:
  - ipotesi: l'attività viene posseduta sul periodo $[t, T]$
  - eventuali flussi generati dal possesso dell'attività sono reinvestiti/finanziati nell'attività stessa
  - $T - t = 1$: 1 anno
  - $T - t = 1/365$: 1 giorno
  - $T - t = 1/250$: 1 giorno (contando solo i giorni in cui i mercati sono aperti)
- ipotesi: $V(u) \geq 0$ per ogni $u > t$ (responsabilità limitata) e $V(t) > 0$ (valore corrente positivo)

PROFITTO/PERDITA

- il profitto/perdita sul periodo $[t, T]$ (holding period P&L) è dato da

$$\text{P\&L} \equiv \text{P\&L}_{t,T} = V(T) - V(t)$$



  variazione di valore dell'attività
- $\text{P\&L}_{t,T}$ rappresenta il guadagno/perdita se si prende una posizione lunga (acquista) il sottostante in t e si liquida la posizione in T
- interpretazione

$$\text{P\&L} > 0 \Rightarrow \text{P\&L} = \text{guadagno}$$

$$\text{P\&L} < 0 \Rightarrow -\text{P\&L} = \text{perdita}$$

- La perdita sul periodo $[t, T]$ è semplicemente

$$L \equiv L_{t,T} = -\text{P\&L}_{t,T} = V(t) - V(T).$$

PROFITTO/PERDITA

- il rendimento viene usualmente misurato su base
  - semplice o composto
  - periodale o annuo
- Rendimento periodale semplice (MISURATO CON UNITÀ DI TEMPO IL PERIODO (t,T))

$$I \equiv I_{t,T} = \frac{\text{P\&L}_{t,T}}{V(t)} = \frac{V(T)}{V(t)} - 1$$

cioè

$$V(T) = V(t)(1 + I_{t,T})$$

- su base annua, il rendimento equivalente è

$$I'_{t,T} = \frac{I_{t,T}}{T-t} = \frac{1}{T-t}\left(\frac{V(T)}{V(t)} - 1\right)$$

$I'_{t,T}$ = RENDIMENTO COSTANTE SU (t,T), IN REGIME D'INTERESSE SEMPLICE

- riesce $-1 \leq I_{t,T} < +\infty$

PROFITTO/PERDITA

▷ Rendimento periodale composto

$$R \equiv R_{t,T} = \log \frac{V(T)}{V(t)}$$

cioè

$$V(T) = V(t)\mathrm{e}^{R_{t,T}}$$

⋆ su base annua, il rendimento equivalente è

$$R'_{t,T} = \frac{R_{t,T}}{T-t} = \frac{1}{T-t} \log \frac{V(T)}{V(t)}$$

$R'_{t,T}$ = = RENDIMENTO (INTENSITÀ) COSTANTE SU (t,T), IN REGIME D'INTERESSE COMPOSTO

⋆ riesce $-\infty \leq R_{t,T} < +\infty$

Profitto/perdita

- Relazioni tra rendimenti semplici e composti

$$I = \mathrm{e}^{R} - 1$$

$$R = \log(1 + I)$$

- se $|R|$ è 'piccolo' (tipicamente quando $T - t$ è piccolo) allora

$$R \approx I$$

$$I = e^{R} - 1 \approx 1 + R - 1$$
$$\approx R$$

e viceversa (formula di Taylor)

$$R = \log(1+I) \approx 0 + I$$
$$\approx I$$

- ESEMPIO: $R = 0.1\% \Rightarrow I = 0.09995\%$
  $R = 1\% \Rightarrow I = 0.99503\%$
  $R = 10\% \Rightarrow I = 9.53102\%$
  (vedi confronto in slide 336)

VaR per perdite e per rendimenti

- relazione tra Value-at-Risk in termini monetari:

$$\mathrm{VaR}_\alpha(L) = \inf\{x \in \mathbb{R} : \mathrm{Prob}[-\mathrm{P\&L} \le x] \ge \alpha\} = F^{-1}_{-\mathrm{P\&L}}(\alpha)$$

e Value-at-Risk in termini di rendimento (composto) $R_{t,T} \equiv R$

$$\mathrm{VaR}_\alpha(-R) = \inf\{x \in \mathbb{R} : \mathrm{Prob}[-R \le x] \ge \alpha\} = F^{-1}_{-R}(\alpha)$$

dove $-R = -R_{t,T} = \log \frac{V(t)}{V(T)}$ è la perdita in termini composti

- P&L è una funzione crescente di R e viceversa; si trovano allora le relazioni $-R = -\log\left(1 - \frac{L_{t,T}}{V(t)}\right)$ $L_{t,T} = V(t)\left(1 - e^{R}\right)$

$$\mathrm{VaR}_\alpha(-R) = -\log\left(1 - \frac{\mathrm{VaR}_\alpha(-\mathrm{P\&L})}{V(t)}\right)$$

$$\mathrm{VaR}_\alpha(-\mathrm{P\&L}) = V(t)\left(1 - \mathrm{e}^{-\mathrm{VaR}_\alpha(-R)}\right)$$

- simili relazioni se si utilizzano i rendimenti semplici ✓

Rendimenti composti: proprietà temporali

▷ "Time aggregation of compound returns": il rendimento periodale composto sul periodo (t, T) è la somma dei rendimenti periodali composti sui sotto-periodi $(t_0, t_1), \ldots, (t_{n-1}, t_n)$, con $t_0 = t < t_1 < \ldots < t_{n-1} < t_n = T$

$$R_{t,T} = R_{t_0,t_1} + \ldots + R_{t_{n-1},t_n}$$

- ⋆ rendimento annuo è somma dei rendimenti giornalieri
- ⋆ nel caso di rendimenti semplici il risultato non è vero, ma

$$1 + I_{t,T} = (1 + I_{t_0,t_1}) \cdot \ldots \cdot (1 + I_{t_{n-1},t_n})$$

- ⋆ nel caso di rendimenti composti/annui?

$$\log \frac{V(T)}{V(t)} = \log \frac{V(T)}{V(t_{n-1})} \frac{V(t_{n-1})}{V(t_{n-2})} \cdots \frac{V(t_1)}{V(t_0)} = R_{t_0,t_1} + \ldots + R_{t_{n-1},t_n}$$

Rendimenti semplici e composti

- Modellizzare rendimenti semplici o composti non è equivalente
- è tipico assumere che i rendimenti siano distribuiti normalmente
- visto il range di $R_{t,T}$ e $I_{t,T}$ questa ipotesi è più adatta ai rendimenti composti MA $P(N(\mu,\sigma^2) < -1)$ POTREBBE ESSERE ≈ 0
- se $R_{t,T} \sim N(\mu, \sigma^2)$ allora $I_{t,T} = e^{R_{t,T}} - 1$ si distribuisce come lognormale traslata
- la distribuzione sarà simile se $T - t$ è piccolo, mentre potrà essere molto diversa quando l'ampiezza dell'intervallo aumenta

Tassi di Cambio

- vantaggio dei rendimenti composti: sia $V(t)$ il tasso di cambio f/d (foreign/domestic)
- $V(t)$ = quantità (PREZZO) di moneta domestica per acquistare 1 unità di valuta straniera al tempo t $\rightsquigarrow$ $\frac{1}{V(t)}$ = quantità di moneta straniera per acquistare 1 unità di valuta domestica al tempo t
- rendimento composto per un investitore domestico: (ACQUISTO MONETA STRANIERA IN t, VENDO IN T)

$$R^d_{t,T} = \log \frac{V(T)}{V(t)}$$

- rendimento composto per un investitore straniero:

$$R^f_{t,T} = \log \frac{V(t)}{V(T)} = -R^d_{t,T}$$

Rendimenti composti: proprietà temporali

- conseguenze dell'additività temporale dei rendimenti composti
- normalità: se la distribuzione congiunta di $R_{t_0,t_1}, \ldots, R_{t_{n-1},t_n}$ è normale, allora $R_{t,T}$ è normale
- momenti, non correlazione seriale
  - ⋆ $E[R_{t_{i-1},t_i}] = \mu_i$, $\mathrm{var}[R_{t_{i-1},t_i}] = \sigma_i^2$
  - ⋆ rendimenti su intervalli disgiunti, $R_{t_0,t_1}, \ldots, R_{t_{n-1},t_n}$, sono incorrelati ($\Rightarrow$ indipendenti) NEL CASO NORMALE

$$\mathrm{cov}[R_{t_{i-1},t_i}, R_{t_{j-1},t_j}] = 0 \qquad \text{per } i \neq j$$

- segue che

$$E[R_{t,T}] = \mu_1 + \ldots + \mu_n$$

$$\mathrm{var}[R_{t,T}] = \sigma_1^2 + \ldots + \sigma_n^2$$

- Se i rendimenti sono correlati?

$$VAR(R_{t,T}) = \sum_{i=1}^{n} \sigma_i^2 + \sum_{i \neq j} \mathrm{cov}(R_{t_{i-1},t_i}, R_{t_{j-1},t_j})$$

Rendimenti composti: proprietà temporali

caso usuale

- nel caso di intervalli di ugual ampiezza e distribuzioni stazionarie:
  - ⋆ $t_i - t_{i-1} = \Delta$ per $i = 1, \ldots, n$, cioè $t_i = t + i\Delta$
  - ⋆ $\mu_i = \mu$, $\sigma_i^2 = \sigma^2$
- riesce allora: media e varianza dei rendimenti composti crescono linearmente col tempo
  - ⋆ $E[R_{t,t+n\Delta}] = n\mu$
  - ⋆ $\text{var}[R_{t,t+n\Delta}] = n\sigma^2$
- "standard deviation rule":

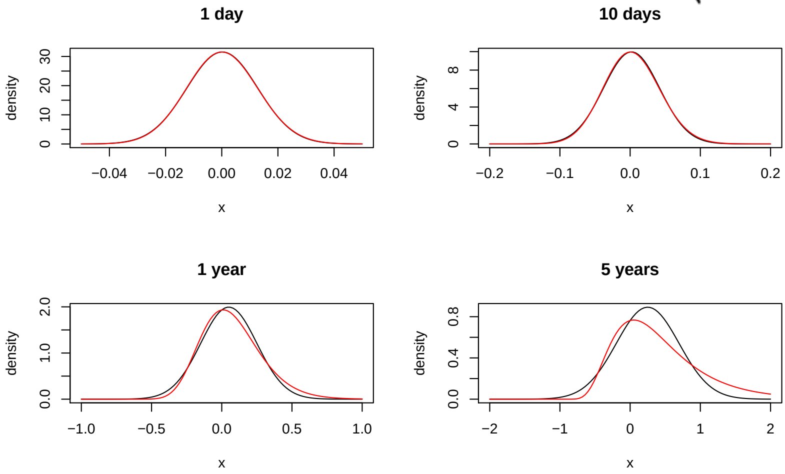
$$\text{sd}[R_{t,t+n\Delta}] = \sigma\sqrt{n}$$

la deviazione standard cresce con la radice del tempo !!

- tale ipotesi può essere testata per verificare la consistenza della non correlazione seriale (e.g. "variance ratio test")
- Nella slide successiva: confronto tra R e I per diversi orizzonti temporali con rendimento atteso 5% e deviazione standard 20% (su base annua)

Rendimenti composti: proprietà temporali

ROSSO : I
NERO : R



P&L DI PORTAFOGLIO

▷ Rendimento di portafoglio: se il portafoglio è composto da N attività con prezzi/valori unitari $V_j(u)$, $j = 1, \ldots, N$ al tempo u e quantità q_j, $j = 1, \ldots, N$, il valore del portafoglio al tempo u è

$$V(u) = \sum_{j=1}^{N} q_j V_j(u)$$

q_J COSTANTE SU (t, T)

- ⋆ ipotesi: la composizione del portafoglio non varia sul periodo (t, T)
- ⋆ profitto/perdita di portafoglio:

$= \sum q_J V_J(T) - \sum q_J V_J(t)$

$$\text{P\&L}_{t,T} = V(T) - V(t) = -L_{t,T} = \sum_{j=1}^{N} q_j \text{P\&L}^{j}_{t,T}$$

- ⋆ rendimento semplice e composto di portafoglio

$$I_{t,T} = \frac{V(T)}{V(t)} - 1, \qquad R_{t,T} = \log \frac{V(T)}{V(t)}$$

P&L DI PORTAFOGLIO

▷ percentuale di ricchezza investita nel titolo j:

RICCHEZZA INVESTITA IN J

$$w_j = \frac{q_j V_j(t)}{V(t)} = \frac{q_j V_j(t)}{\sum_{j=1}^N q_j V_j(t)}$$

RICCHEZZA TOTALE

riesce $\sum_{j=1}^N w_j = 1$

- ⋆ rendimento semplice di portafoglio è la media aritmetica ponderata dei rendimenti semplici ("additivity across assets")

$$\boxed{I_{t,T} = \sum_{j=1}^N w_j I_{t,T}^j}$$

- ⋆ rendimento composto di portafoglio è la media ponderata esponenziale dei rendimenti composti

$$R(t,T) = \log\left(\sum_{j=1}^N w_j \mathrm{e}^{R_{t,T}^j}\right)$$

$e^R = \sum w_j e^{R^j}$

RENDIMENTO DI PORTAFOGLIO

$$I_{t,T} = \frac{V(T)}{V(t)} - 1 = \frac{\sum q_J V_J(T) - \sum q_J V_J(t)}{V(t)}$$

$$= \frac{\sum q_J \left(V_J(T) - V_J(t) \right)}{V(t)}$$

$$= \sum \underbrace{\frac{q_J V_J(t)}{V(t)}}_{w_J} \underbrace{\frac{V_J(T) - V_J(t)}{V_J(t)}}_{I^J_{t,T}}$$

$$= \sum w_J I^J_{t,T}$$

P&L di portafoglio

- Se i rendimenti delle attività componenti il portafoglio sono "'piccole", allora vale l'approssimazione (usare $e^x \approx 1 + x$ e $\log(1 + x) \approx x$ quando $x \to 0$)

$$\boxed{R_{t,T} \approx \sum_{j=1}^{N} w_j R^j_{t,T}}$$

  - ⋆ se la distribuzione congiunta di $R^1_{t,T}, \ldots, R^N_{t,T}$ è normale, in generale non è nota la distribuzione di $R_{t,T}$ (logaritmo di una combinazione lineare di lognormali)
  - ⋆ tuttavia, l'approssimazione sopra consente di assumere normalità delle componenti e del portafoglio allo stesso tempo
  - ⋆ l'approssimazione è valida per orizzonti temporali limitati

RENDIMENTI COMPOSTI DI PORTAFOGLIO

$$R_{t,T} = \log \frac{V(T)}{V(t)} = \log \frac{\sum q_J V_J(T)}{V(t)}$$

$$= \log \sum \underbrace{\frac{q_J V_J(t)}{V(t)}}_{w_J} \underbrace{\frac{V_J(T)}{V_J(t)}}_{e^{R^J_{t,T}}}$$

APPROSSIMAZIONE, SE $R^J_{t,T}$ "PICCOLI"

$\log(1+x) \approx x$

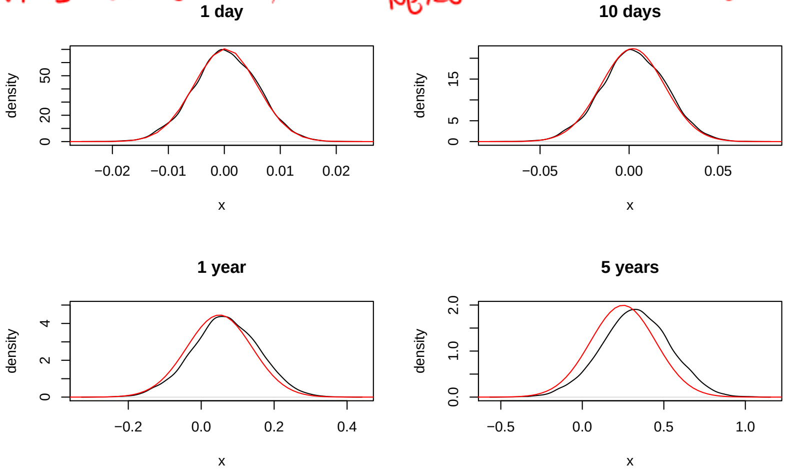
$$R_{t,T} \approx \sum w_J \underbrace{e^{R^J_{t,T}}}_{e^x \approx 1+x} - 1 \approx \sum w_J (1 + R^J_{t,T}) - 1$$

$$= \sum w_J R^J_{t,T}$$

Rendimenti di portafoglio

CONFRONTO TRA $R_{t,T}$ E $\sum w_J R^J_{t,T}$

(VIA SIMULAZIONE) NERO ROSSO



Value-at-Risk

- VaR con distribuzione normale: $(t, T) = (t, t + n\Delta)$, $R_{t_{i-1},t_i} \sim N(\mu, \sigma^2)$, rendimenti incorrelati
  - ⋆ $E[R_{t,t+n\Delta}] = n\mu$, $\mathrm{sd}[R_{t,t+n\Delta}] = \sigma\sqrt{n}$ ← " SQ RULE "
  - ⋆ il Value-at-Risk sull'orizzonte temporale $(t, t + n\Delta)$ è

$$\mathrm{VaR}_\alpha(-R_{t,t+n\Delta}) = -\mu n + \sigma\sqrt{n}\Phi^{-1}(\alpha)$$

Se $\mu = 0$, allora

$$\boxed{\mathrm{VaR}_\alpha(-R_{t,t+n\Delta}) = \sqrt{n}\,\mathrm{VaR}_\alpha(-R_{t,t+\Delta})}$$

⇝ utile per stimare i parametri su intervalli più piccoli (più dati) e estendere a intervalli più ampi

  - ⋆ andamento rispetto all'orizzonte temporale n (se $\alpha > 50\%$)?

$$\mathrm{VaR} \uparrow n \quad \text{se} \quad \mu \leq 0$$

$$\mathrm{VaR} \quad \text{prima crescente, poi decrescente, se} \quad \mu > 0$$

VaR IN FUNZIONE DI n

(N° PERIODI)

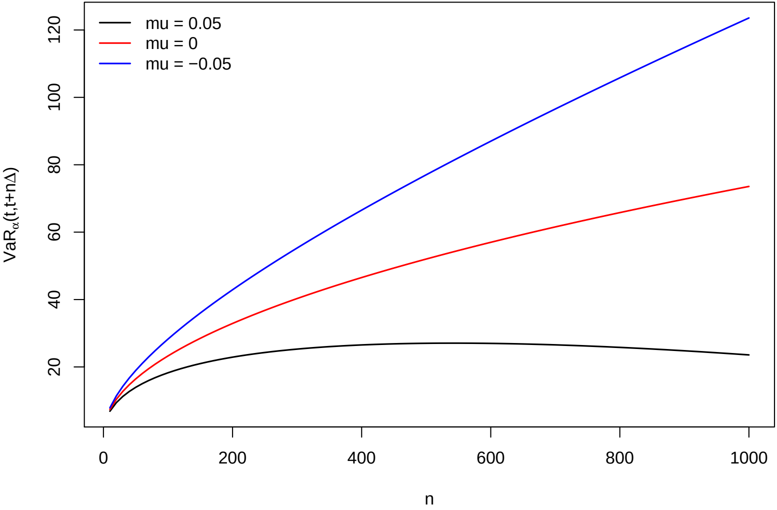
$$\text{VaR}_\alpha(-R_{t,t+nA}) = -\mu n + \sigma\sqrt{n}\,\underbrace{\Phi^{-1}(\alpha)}_{\geq 0 \text{ SE } \alpha \geq 0.5}$$

* SE $\mu \leq 0$ $\quad \text{VaR}_\alpha \nearrow$ CON n

* SE $\mu > 0$ $\quad \dfrac{\partial \text{VaR}_\alpha}{\partial n} = -\mu + \dfrac{\sigma\,\Phi^{-1}(\alpha)}{2\sqrt{n}} \begin{array}{l} > 0 \\ < 0 \end{array}$

$\text{VaR}_\alpha \nearrow \searrow$ CON n

Value-at-Risk



Value-at-Risk: approccio parametrico

O CMQ UN NUMERO LIMITATO

- ▷ vantaggi
  - ⋆ solo due (o tre) parametri da stimare
  - ⋆ si estende a periodi di lunghezza n via la regola della radice quadrata (nel caso normale)
  - ⋆ si estende al caso di portafogli
- ▷ svantaggi
  - ⋆ scelta di un modello che consenta code pesanti e asimmetria? QUALE SCEGLIERE ?
  - ⋆ rischio di parametro: il metodo richiede la stima dei parametri μ, σ (e ν nel caso t di Student) $\Rightarrow$ intervalli di confidenza per il Value-at-Risk INCERTEZZA SUI PARAMETRI => INCERTEZZA SUL VaR
  - ⋆ rischio di modello: il metodo richiede la scelta di un modello $\Rightarrow$ ogni modello è sbagliato

Value-at-Risk: historical simulation

- Per calcolare il VaR con ragionevole precisione, la lunghezza m del campione deve essere sufficientemente grande; ad esempio, per $\alpha = 99\%$ non ha senso usare meno di $m = 100$
- tipicamente si usa una finestra includente gli m rendimenti più recenti; scelte tipiche per rendimenti giornalieri
  - $m = 250$ (un anno)
  - $m = 1000$ (quattro anni)
- La scelta di m influenza il VaR
- m troppo elevato rende il VaR insensibile avendo ogni osservazione un peso basso
- al muoversi della finestra, il VaR ha dei movimenti improvvisi (salti) dovuti all'inclusione di nuove osservazioni/esclusione di vecchie osservazioni

ROBUSTEZ-
ZA DEL
VaR

Value-at-Risk: historical simulation

- ESEMPIO: $l_1 = 1\%$, $l_2 = -2\%$, $l_3 = 0\%$, $l_4 = -1\%$, $l_5 = 2.5\%$, $l_6 = -1\%$, $l_7 = 3\%$, $l_8 = 0.5\%$, $l_9 = 1\%$, $l_{10} = 4\%$
  - statistica d'ordine:

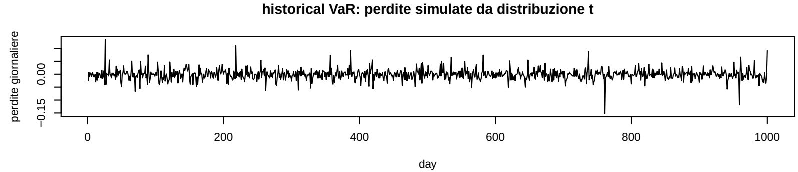
$$l_{(1)} = -2\%,\ l_{(2)} = l_{(3)} = -1\%,\ l_{(4)} = 0\%,\ l_{(5)} = 0.5\%,$$
$$l_{(6)} = l_{(7)} = 1\%,\ l_{(8)} = 2.5\%,\ l_{(9)} = 3\%,\ l_{(10)} = 4\%$$

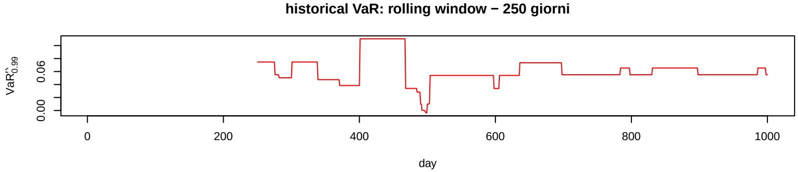
    si trova $\mathrm{VaR}_{0.9} = l_{(9)} = 3\%$
  - se $l_{10} = 10\%$ invece di 4%, il VaR non cambia $\rightsquigarrow$ il VaR è "blind to the tail", non dà informazioni sulle perdite superiori al VaR
  - Se si usa una finestra di 10 osservazioni e l_{11} è la nuova perdita, mentre esce $l_1 = 1\%$ dal campione: se $l_{11} \le 3\%$ allora $\mathrm{VaR}_{0.9} = 3\%$; se $3\% < l_{11} \le 4\%$ allora $\mathrm{VaR}_{0.9} = l_{11}$; se $l_{11} > 4\%$ allora $\mathrm{VaR}_{0.9} = 4\%$

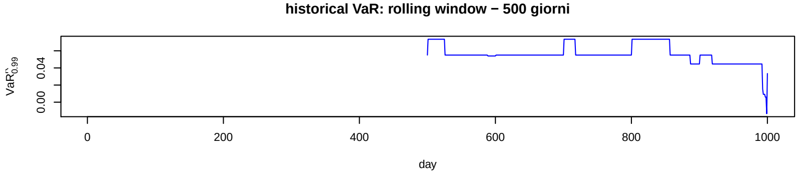
VaR DISCONTINUO RISPETTO A l_{11}!

$$\mathrm{VaR}^{\text{NUOVO}}_{90\%}(L) = \begin{cases} 3\% & l_{11} \le 3\% \\ l_{11} & 3\% < l_{11} \le 4\% \\ 4\% & l_{11} > 4\% \end{cases}$$

Value-at-Risk: historical simulation







Weighted Historical Value-at-Risk (WHVaR)

PESO : 1 1 ---- 1
l_1 l_2 ---- l_m
1 2 ---- m
VaR

- Per rimediare alla mancanza di robustezza/eccessiva sensibilità dell'historical VaR, si usa uno schema di pesi
  - ⋆ nel VaR, tutte le osservazioni hanno lo stesso peso
  - ⋆ nel WHVaR, si usano pesi esponenzialmente decrescenti a seconda di quanto recente è l'osservazione: date le perdite $l_1, \ldots, l_m$, il peso di l_i è

$$w_i = \frac{\eta^{m-i}(1-\eta)}{1-\eta^m} \propto \eta^{m-i}$$

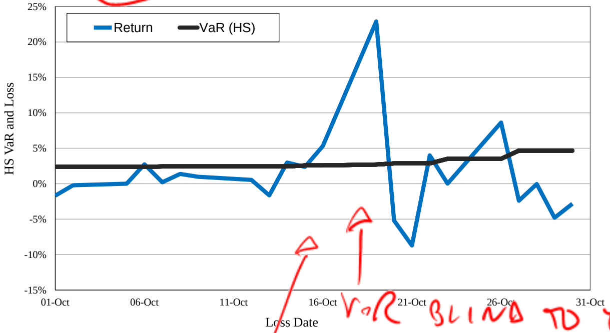
    con $0 < \eta < 1$ (tipicamente $\eta > 0.9$)
  - ⋆ in questo modo l'osservazione più vecchia che esce dal campione ha poco peso, mentre ha peso massimo la più recente
  - ⋆ il ruolo di m è meno importante
  - ⋆ si calcola poi il percentile della distribuzione (tenendo conto dei pesi)

PESO : η^{m-1} η^{m-2} ---- η 1
l_1 l_2 ---- l_{m-1} l_m
1 2 ---- $m-1$ m
WH-VaR

Weighted Historical Value-at-Risk

Figure 2.2 A:
Historical Simulation VaR and Daily Losses from
Long S&P500 Position, October 1987 → "BLACK MONDAY"

S&P 500 −22% IN UN GIORNO



VaR BLIND TO THE TAIL

Elements of Financial Risk Management Second Edition © 2012 by Peter Christoffersen

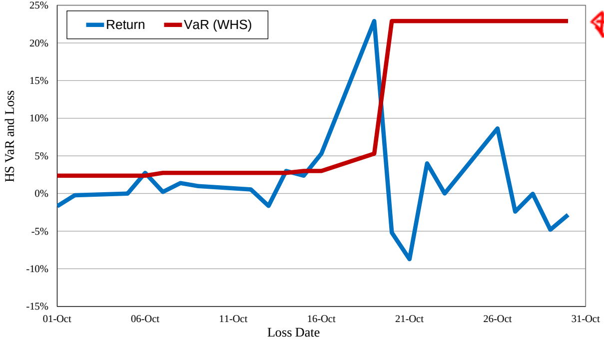
"GHOST FEATURE"

"$\ell = 22\%$" ENTRA NEL CAMPIONE

Weighted Historical Value-at-Risk

16

Figure 2.2 B:
Historical Simulation VaR and Daily Losses from
Long S&P500 Position, October 1987



REAGISCE MEGLIO, DÀ PESO ALL'ULTIMA OSSERVAZIONE

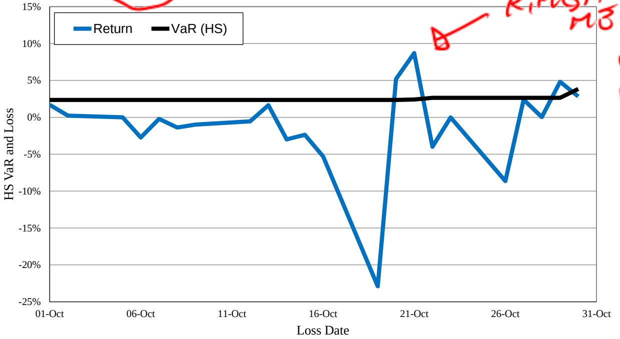
Elements of Financial Risk Management Second Edition © 2012 by Peter Christoffersen

Weighted Historical Value-at-Risk

17

Figure 2.3 A:
Historical Simulation VaR and Daily Losses from
Short S&P500 Position, October 1987

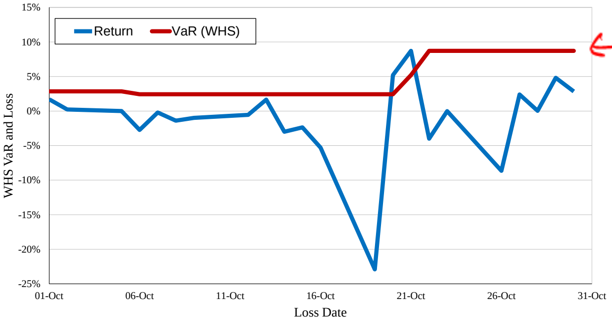
RIMBALZO DEL MERCATO, VaR NON REAGISCE



Elements of Financial Risk Management Second Edition © 2012 by Peter Christoffersen

Weighted Historical Value-at-Risk

Figure 2.3 B:
Weighted Historical Simulation VaR and Daily Losses
from **Short** S&P500 Position, October 1987



Elements of Financial Risk Management Second Edition © 2012 by Peter Christoffersen

Value-at-Risk: bootstrap

- problemi della simulazione storica:
  - a differenza del VaR parametrico, non si può estendere facilmente il caso uniperiodale a quello n-periodale
  - si potrebbe applicare il metodo direttamente a periodi di lunghezza n al costo di una perdita notevole di dati (e.g. con 250 rendimenti giornalieri ci sono solo 12 rendimenti mensili)
- alternativa: bootstrap non parametrico (Efron (1979))
  - tecnica introdotta in statistica inferenziale per sfruttare campioni limitati
  - I) idea: (ri)estrarre con reinserimento un campione da quello dato, di uguale lunghezza
  - II) calcolare la statistica di interesse sul nuovo campione VaR, ES, ...
  - III) ripetere la procedura un numero elevato di volte
  - IV) si ottiene così la distribuzione campionaria della statistica di interesse ⇒ media, intervalli di confidenza, ...

VaR, ES

Value-at-Risk: bootstrap

▷ bootstrap VaR, 1 periodo

1. è dato un campione $r_1, \ldots, r_m$ di rendimenti uniperiodali (e.g. giornalieri)

I) 2. si costruisce un nuovo campione di medesima lunghezza m via estrazione con rimpiazzo, indicato con $\widehat{r}_{1,1}, \ldots, \widehat{r}_{1,m}$

II) 3. si calcola (via simulazione storica) il Value-at-Risk dal nuovo campione, indicato con $\widehat{\mathrm{VaR}}_{\alpha,1}(-R)$

III) 4. si ripetono 2. e 3. un numero scelto M di volte, ottenendo M campioni di lunghezza m

5. da ognuno degli M campioni si estrae il Value-at-Risk, indicati con $\widehat{\mathrm{VaR}}_{\alpha,1}(-R), \ldots, \widehat{\mathrm{VaR}}_{\alpha,M}(-R)$

IV) 6. si calcola poi la stima di bootstrap del VaR via

$$\widehat{\mathrm{VaR}}_{\alpha,\mathrm{bootstrap}}(-R) = \frac{1}{M} \sum_{i=1}^{M} \widehat{\mathrm{VaR}}_{\alpha,i}(-R)$$

VALUE-AT-RISK: BOOTSTRAP

- bootstrap VaR, 1 periodo
  - si può usare il campione di bootstrap dei Value-at-Risk per costruire intervalli di confidenza per $\widehat{\mathrm{VaR}}_{\alpha,\mathrm{bootstrap}}(-R)$
  - si calcolano i Value-at-Risk per ogni campione bootstrap come in 1.-5., $\widehat{\mathrm{VaR}}_{\alpha,1}(-R), \ldots, \widehat{\mathrm{VaR}}_{\alpha,M}(-R)$ NECESSARIAMENTE
  - si sceglie un livello di confidenza α', non collegato con il livello di confidenza del Value-at-Risk α!
  - si calcolano poi il $(1-\alpha')/2$ e il $(1+\alpha')/2$ quantili della distribuzione $\widehat{\mathrm{VaR}}_{\alpha,1}(-R), \ldots, \widehat{\mathrm{VaR}}_{\alpha,M}(-R)$

IV)

- il metodo funziona bene se il campione iniziale non è troppo piccolo e se proviene da una distribuzione ragionevolmente simmetrica

IDEA: LA DISTRIBUZIONE EMPIRICA È "VICINA" A QUELLA "CORRETTA" => SIMULO DALLA DISTRIBUZIONE EMPIRICA

Value-at-Risk: bootstrap

- osservazioni sul metodo bootstrap
  - ogni ricampionamento del campione originale ne genera uno potenzialmente diverso
  - il Value-at-Risk calcolato su ogni ricampionamento sarà (potenzialmente) diverso da quello del campione originale

    DEVE ESSERE COSÌ
  - in ogni ricampionamento, una data osservazione del campione originale potrà apparire più volte, o non apparire per niente
  - una osservazione che appare ripetuta nel campione originale sarà ricampionata con probabilità proporzionale alla sua frequenza
  - ~~l'effetto di questa distorsione — i VaR di ogni bootstrap differiscono da quello empirico calcolato sul campione originale — se la dimensione campionaria m è sufficientemente grande~~
- idea del metodo bootstrap: la variabilità delle stime di boostrap intorno alla stima empirica è simile/mima la variabilità della stima empirica intorno al valore "vero"
- per campioni di dimensione m ragionevolmente elevata, la distribuzione empirica è vicina alla popolazione $\Rightarrow$ il bootstrap (campionare con rimpiazzo dalla distribuzione empirica) non differisce molto dal campionamento casuale dalla popolazione

Value-at-Risk: bootstrap PER MULTIPERIODO

- bootstrap VaR, n periodi: $(t, T) = (t, t + n\Delta)$
  1. è dato un campione $r_1, \ldots, r_m$ di rendimenti uniperiodali (e.g. giornalieri)

I) 2. si costruisce un nuovo campione di lunghezza n via estrazione con rimpiazzo, indicato con $\widehat{r}_{1,1}(1), \ldots, \widehat{r}_{1,1}(n)$

I') 3. si ottiene un rendimento sul periodo $(t, t + n\Delta)$ via

$$\widehat{r}_{1,1}(t, t + n\Delta) = \sum_{j=1}^{n} \widehat{r}_{1,1}(j)$$

"RENDIMENTO PERIODALE = SOMMA RENDIMENTI UNIPERIODALI"

I'') 4. si ripetono 2.-3. m volte in maniera da ottenere un campione $\widehat{r}_{1,1}(t, t + n\Delta), \ldots, \widehat{r}_{1,m}(t, t + n\Delta)$ di rendimenti n-periodali

II) 5. si calcola (via simulazione storica) il Value-at-Risk dal campione multi periodale, indicato con $\widehat{\mathrm{VaR}}_{\alpha,1}(-R)$

III) 6. si ripetono 2.-4. un numero scelto M di volte, ottenendo M stime di Value-at-Risk, indicati con $\widehat{\mathrm{VaR}}_{\alpha,1}(-R), \ldots, \widehat{\mathrm{VaR}}_{\alpha,M}(-R)$

IV) 7. si calcola poi la stima di bootstrap del VaR via

$$\widehat{\mathrm{VaR}}_{\alpha,\mathrm{bootstrap}}(-R) = \frac{1}{M} \sum_{i=1}^{M} \widehat{\mathrm{VaR}}_{\alpha,i}(-R)$$

Value-at-Risk: bootstrap

▷ bootstrap VaR, n periodi: $(t, T) = (t, t + n\Delta)$

IV)

- ⋆ similmente si aggiungono intervalli di confidenza ✓
- ⋆ l'ipotesi fondamentale è che i rendimenti siano indipendenti su periodi diversi $\Rightarrow$ estrazioni indipendenti dallo stesso campione

È UGUAL-
MENTE
DISTRIBU-
ITI

- ⋆ lo schema di base può essere modificato per includere correlazione seriale dei rendimenti e volatilità stocastica (e.g. tramite process GARCH = Generalized Autoregressive Conditional Heteroskedasticity)

VALUE-AT-RISK: BOOTSTRAP

- vantaggi:
  - procedura non parametrica (model-free) $\Rightarrow$ non richiede stima di parametri e scelta di un modello
  - procedura semplice da implementare
  - implicitamente include asimmetria e code pesanti SE DISPONIBILI...
  - si estende in maniera relativamente semplice al caso di autocorrelazione e/o volatilità stocastica
- svantaggi:
  - molto sensibile alla lunghezza del data set: tipicamente si usano da 250 a 1000 osservazioni giornaliere (da 1 a 4 anni); per ottenere la stessa accuratezza del metodo parametrico sono richiesti campioni di dimensioni anche superiori
  - $\Rightarrow$ serie temporali potrebbero non essere disponibili
  - lunghe serie sono richieste per includere eventi estremi
  - "Ghost feature": il metodo reagisce lentamente a variazioni nel rischio di mercato
  - nel caso n-periodale la procedura può essere computazionalmente pesante

Teoria dei Valori Estremi

- la teoria dei valori estremi (EVT) è una branca della probabilità e statistica il cui scopo è
  - costruire modelli per la misurazione (ie calcolare probabilità, quantili, momenti, ...) di eventi estremi
  - sviluppare procedure per la stima di tali modelli
- cosa significa estremo? alcune possibili definizioni:
  - *eccedente l'ordinario, usuale, o atteso*
  - *non frequente, non comune, raro*
  - *remoto in ogni direzione*

TEORIA DEI VALORI ESTREMI

- alcune citazioni famose sulla Teoria dei Valori Estremi (fonte: `http://www.isse.ucar.edu/extremevalues/quotes.html`)
  - Sir Ronald Fisher:
    "The 'one chance in a million' will occur, with no less and no more than its appropriate frequency, however surprised we may be that it should occur to us."
  - Emil Gumbel:
    "Il est impossible que l'improbable n'arrive jamais."
    "Il y aura toujours une valeur qui dépassera toutes les autres."
  - John Tukey:
    "As I am sure almost every geophysicist knows, distributions of actual errors and fluctuations have much more straggling extreme values than would correspond to the magic bell-shaped distribution of Gauss and Laplace."
    ≡ NORMALE

Teoria dei Valori Estremi

- formalizzando i concetti visti prima, si consideriuna sequenza di variabili aleatorie
$$X_1, X_2, \ldots, X_n, \ldots$$
independenti e identicamente distribuiti (iid)
- tipica interpretazione: l'indice n rappresenta il tempo, cioè X_n potrebbe essere
  - ⋆ livello del mare (in una data posizione) / ammontare nevicata / ammontare pioggia / velocità del vento / rendimento / ... al giorno n-esimo
  - ⋆ oppure minuto / ora / settimana / mese / anno
- n potrebbe essere l'indice dell'n-esima osservazione (eg X_n = magnitudo del n-esima scossa di terremoto) NON EQUI INTERVALLATE
- interpretazione "spaziale" di n: X_n potrebbe essere
  - ⋆ livello del mare all'n-esima posizione
  - ⋆ sinistro registrato dall'n-esima polizza assicurativa
  - ⋆ tempo registrato dall'n-esimo atleta in una gara

Teoria dei Valori Estremi

- Approccio classico alla EVT (block-maxima approach) permette di approssimare probabilità tipo

$$P[\max\{X_1, \dots, X_n\} > x] \approx ?$$

per n grande, cioè la probabilità che almeno una osservazione ecceda il livello x

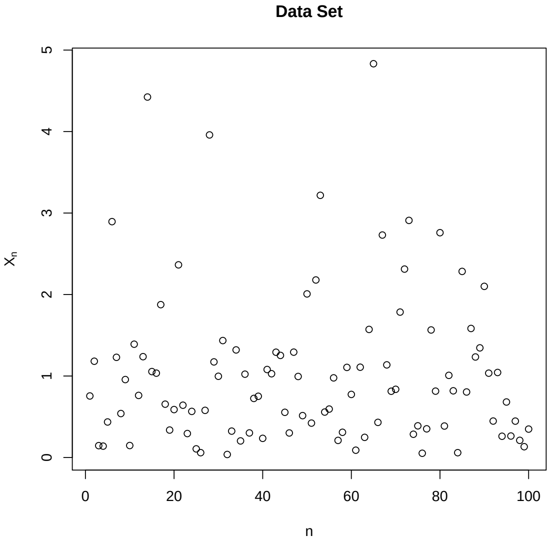
- si possono considerare minimi, dal momento che

$$\min\{X_1, \dots, X_n\} = -\max\{-X_1, \dots, -X_n\}$$

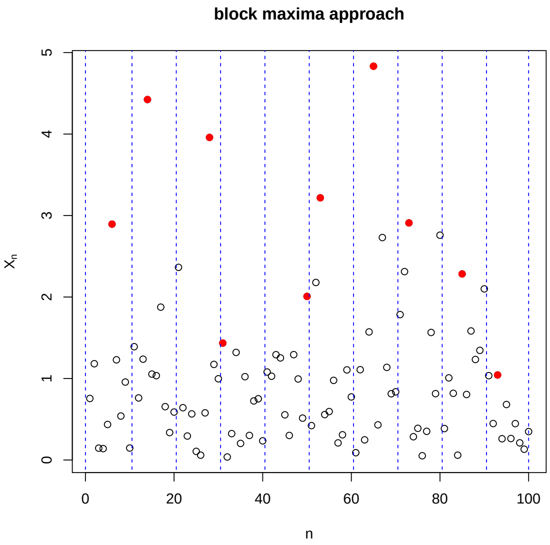
"OGNI MINIMO È UN MASSIMO"

- in questo approccio,
  - un evento estremo corrisponde alla massima osservazione
  - il campione viene diviso in blocchi e il massimo viene estratto da ogni blocco ⇝ la scelta del blocco è critica
  - valori inferiori al massimo vengono scartati

Teoria dei Valori Estremi



Teoria dei Valori Estremi

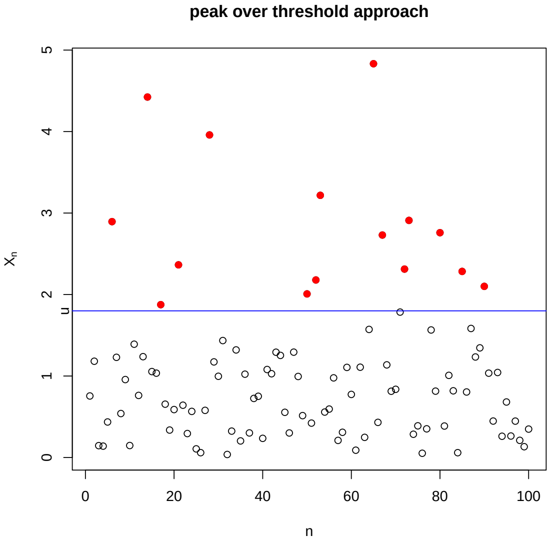


Teoria dei Valori Estremi

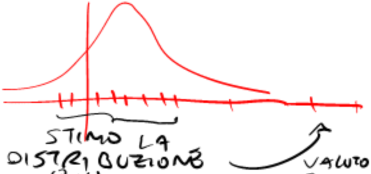
- Un approccio alternativo è chiamato Peak over Threshold (POT) approach: viene considerata una soglia elevata u e si possono approssimare quantità tipo $\approx$?
  - ⋆ $P[X_n > u + y | X_n > u]$, $y \geq 0$
  - ⋆ $E[X_n - u | X_n > u]$

  per n fissato e u grande
- nell'approccio POT
  - ⋆ un evento estremo corrisponde a un eccesso della soglia u
  - ⋆ valori inferiori a u vengono scartati
  - ⋆ la scelta di u è critica
  - ⋆ si possono considerare eccessi sotto una soglia u
    $P[X_n \leq u - y | X_n \leq u]$, $y \geq 0$

Teoria dei Valori Estremi



Teoria dei Valori Estremi



▷ Vantaggi dell'EVT

- ⋆ permette di utilizzare al meglio i dati disponibili per quantificare eventi estremi
- ⋆ tecniche statistiche standard non sono adatte:
  - il problema in questione richiede di valutare probabilità di eventi nella coda della distribuzione, possibilmente al di là del range dei dati
  - la teoria standard tratta i valori estremi come outliers
  - errori di stima possono essere amplificati quando si valutano eventi estremi
- ⋆ l'EVT invece usa una teoria asintotica (come nel CLT) che permette di calcolare le quantità richieste senza far ipotesi sulla distribuzione sottostante dei dati ⇝ l'errore di modello è meno importante
- ⋆ l'EVT fornisce una chiara indicazione sulla natura della coda della distribuzione dei dati

Teoria dei Valori Estremi

- ▷ difetti della EVT
  - ⋆ teoria asintotica ⇝ richiede larghezza campionaria sufficientemente grande perchè l'approssimazione sia efficace ⇝ trade-off tra
    - bias (distorsione del modello)
    - varianza (precisione degli stimatori)
  - ⋆ EVT di base assume osservazioni iid ⇝ spesso non soddisfatta
- ▷ estensioni della EVT
  - ⋆ EVT può essere estesa a osservazioni stazionarie (dipendenti)
  - ⋆ tuttavia, i dati spesso esibiscono trend e/o stagionalità ⇝ non stazionarie
  - ⋆ studio congiunto di due o più fenomeni correlati, eg
    - livello del mare (velocità del vento, quantità di pioggis, neve . . .) in due località diverse
    - velocità del vento e quantità di pioggia in una data località
    - serie finanziarie diverse — FX e azionario

  ⇝ EVT multivariata

TEORIA DEI VALORI ESTREMI

- Sia $X_1, \dots, X_n, \dots$ una successione di variabili aleatorie
- siamo interessati in valutazioni asintotiche sul massimo campionario e il minimo campionario

$$m_n = \min\{X_1, \dots, X_n\}, \; M_n = \max\{X_1, \dots, X_n\}$$

- si può estendere la considerazione alle k osservazioni più grandi

$$X_{(1)} \qquad X_{(2)} \qquad\qquad X_{(n-1)} \qquad X_{(n)}$$

$$X_{n:n} \leq X_{n-1:n} \leq \dots \leq X_{2:n} \leq X_{1:n},$$

dove $X_{n:n} = m_n$ è il minimo, $X_{n-1:n}$ è il secondo più piccolo, $\dots$, $X_{2:n}$ è il secondo più grande e $X_{1:n} = M_n$ è il massimo

Teoria dei Valori Estremi

- l'ipotesi di base è che le variabili aleatorie $X_1, \dots, X_n, \dots$ siano independenti e identicamente distribuite (iid); indichiamo con F la funzione di ripartizione comune agli X_i's
- è immediato ricavare la legge di M_n, F_{M_n}:

$$F_{M_n}(x) = F(x)^n$$

- Il comportamento asintotico di M_n quando $n \to +\infty$ è facilmente descritto: M_n converge in distribuzione alla variabile aleatoria degenere in $\bar{x} := \inf\{x \in \mathbb{R} : F(x) = 1\}$:

$$M_n \to^{\mathrm{d}} \bar{x}$$

$$\forall x \quad F_{M_n}(x) \to F_{\bar{x}}(x) = \begin{cases} 1 & x \geq \bar{x} \\ 0 & x < \bar{x} \end{cases}$$

IN OGNI PUNTO DI CONTINUITÀ DI $F_{\bar{x}}$

- risultato non utile, è necessaria una normalizzazione come nel CLT

$$F_{M_n}(x) = P(M_n \le x)$$

$$= P(\underbrace{\max\{X_1, \ldots, X_n\} \le x}_{=}\,)$$

$$= P(X_1 \le x, \ldots, X_n \le x)$$

INDIP.

$$= P(X_1 \le x) \cdot \ldots \cdot P(X_n \le x)$$

$$= F(x)^n$$

UGUAL DISTRIBUZIONE

SE $F(x) < 1$ $\Rightarrow \lim_{n \to +\infty} F_{M_n}(x) = 0$

SE $F(x) = 1$ $\Rightarrow \lim_{n \to +\infty} F_{M_n}(x) = 1$

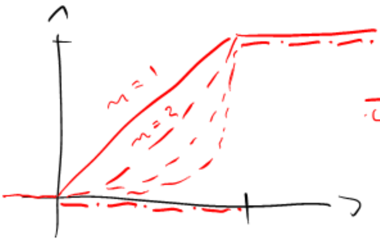
$$\lim_{n \to +\infty} F_{M_n}(x) = \begin{cases} 0 & x < \bar{x} = \operatorname{ess\,sup} X_i \\ 1 & x > \bar{x} \end{cases}$$

ESEMPIO: $F(x) = x \quad 0 < x < 1$

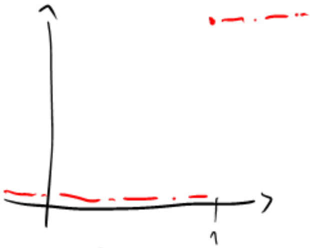
$$X_i \sim U(0,1)$$

$$F_{M_n}(x) = \begin{cases} 0 & x < 0 \\ x^n & 0 \le x \le 1 \\ 1 & x > 1 \end{cases} \leftarrow \bar{X} = 1$$

$\Downarrow$



LIMITE



DEGENERE

LEGGE FORTE DEI GRANDI NUMERI

E TEOREMA LIMITE CENTRALE

$X_1 \ldots X_n \ldots$ IID $\quad E(|X_i|) < +\infty$

$S_n = X_1 + \ldots + X_n$

$\bar{X}_n = S_n / n$ "MEDIA EMPIRICA"

$\bar{X}_n \xrightarrow[n \to \infty]{} E[X_1]$ CON PROB. 1

DEGENERE

NORMALIZZAZIONE RICHIESTA AL FINE DI OTTENERE UN LIMITE NON DEGENERE! SE $E(X_1) = \mu$

$VAR(X_1) = \sigma^2$

$$\frac{\bar{X}_n - E(\bar{X}_n)}{SD(\bar{X}_n)} = \underbrace{\frac{\bar{X}_n - \mu}{\sigma / \sqrt{n}}} \xrightarrow[n \to \infty]{} N(0,1)$$

$$\underbrace{a_n \bar{X}_n + b_n}$$

CAMBIO DI LOCAZIONE E SCALA!

È POSSIBILE APPLICARE LA STESSA IDEA AL MASSIMO M_n?

$\exists\, a_n, b_n : a_n M_n + b_n \longrightarrow ??$

Teoria dei Valori Estremi

▷ Teorema dei Tre Tipi (Tippet-Fisher 1928, Gnedenko 1943).
Siano $X_1, \ldots, X_n, \ldots$ sono variabili aleatorie iid e $M_n = \max\{X_1, \ldots, X_n\}$; se esistono $a_n > 0$, $b_n \in \mathbb{R}$ e una variabile aleatoria L non degenere con funzione di ripartizione H tale che

$$a_n M_n + b_n \to^{\mathrm{d}} L.$$

$F_{a_n M_n + b_n}(x) \to H(x)$ $\forall x$: H è continua in x

Allora, a meno di un cambio di locazione e scala, H deve essere una di

- ⋆ Tipo I - Gumbel
- ⋆ Tipo II - Fréchet
- ⋆ Tipo III - Weibull negativa

Note come distribuzioni dei valori estremi

Teoria dei Valori Estremi

- le distribuzioni dei valori estremi sono
  - Tipo I - Gumbel

$$\boxed{H^{\mathrm{I}}(x) = \mathrm{e}^{-\mathrm{e}^{-x}}, \ -\infty < x < \infty}$$

SUPPORTO

$\mathbb{R}$

  - Tipo II - Fréchet

$$\boxed{H^{\mathrm{II}}_{\alpha}(x) = \begin{cases} 0 & \text{se } x \leq 0 \\ \mathrm{e}^{-x^{-\alpha}} & \text{se } x > 0 \end{cases}}$$

$(0, +\infty)$

con $\alpha > 0$

  - Tipo III - Weibull negativa

SE X HA DISTRIBUZIONE WEIBULL $\Rightarrow$ $-X$ HA DISTRIBUZIONE

$$\boxed{H^{\mathrm{III}}_{\alpha}(x) = \begin{cases} \mathrm{e}^{-(-x)^{\alpha}} & \text{se } x < 0 \\ 1 & \text{se } x \geq 0 \end{cases}}$$

$(-\infty, 0)$

con $\alpha > 0$

Teoria dei Valori Estremi

- per ottenere il cambio di locazione e scala basta sostituire x con $\frac{x-\mu}{\sigma}$, con $\sigma > 0$ e $\mu \in \mathbb{R}$, e cambiando il supporto di conseguenza
- i tre tipi si possono riassumere con la singola espressione, nota come distribuzione generalizzata dei valori estremi (GEV)

$$G_\xi(x) = \exp\left(-\left[1+\xi x\right]^{-1/\xi}\right),\ 1+\xi x > 0$$

$\frac{x-\mu}{\sigma}$

  - ⋆ $\xi \to 0$ si trova la Gumbel
  - ⋆ $\xi = 1/\alpha > 0$ si trova la Frèchet
  - ⋆ $\xi = -1/\alpha < 0$ si trova la Weibull negativa

- il Teorema dei Tre Tipi ci dice che, per n grande, riesce per qualche $\xi \in \mathbb{R}$,

$$\boxed{P(M_n \leq x) \approx G_\xi\left(\frac{x-\mu}{\sigma}\right) \quad \text{per ogni } x \in \mathbb{R}}$$

(μ, σ) + $M_n \sim GEV$ PER n GRANDE

$\boxed{\xi = 0}$

$$\lim_{\xi \to 0} \exp\left(- (1+\xi x)^{-1/\xi} \right) = \qquad \xi = \frac{1}{\alpha}$$

$$= \lim_{\alpha \to \infty} \exp\left(- \Big(\underbrace{\left(1 + \frac{x}{\alpha}\right)^{\alpha}}_{\to e^{x}} \Big)^{-1} \right)$$

$$= e^{-e^{-x}} , \qquad x \in \mathbb{R}$$

$1 + \xi x > 0$ SEMPRE VERA SE $\xi \to 0$

$\boxed{\xi > 0}$ $G_\xi(x) = \exp\left(-(1+\xi x)^{-1/\xi}\right)$ $\xi = \frac{1}{\alpha}$

$= \exp\left(-\left(1+\frac{x}{\alpha}\right)^{-\alpha}\right)$

$= \exp\left(-\left(\frac{x+\alpha}{\alpha}\right)^{-\alpha}\right)$ FRECHET

$= H_\alpha^{I}\left(\frac{x+\alpha}{\alpha}\right)$ $\mu = -\alpha$, $\sigma = \alpha$

$1+\xi x > 0 \iff x > -\frac{1}{\xi} = -\alpha$

APPROSSIMAZIONE

$$a_n M_n + b_n \xrightarrow{d} L$$

$$F_{a_n M_n + b_n}(x) \longrightarrow H(x) \qquad \forall x$$

$$P(a_n M_n + b_n \leq x) \approx H(x) \qquad \forall x,$$
n GRANDE

$$P(M_n \leq x) = P(a_n M_n + b_n \leq a_n x + b_n)$$

$$\approx H(\underbrace{a_n x + b_n}_{\text{CAMBIO DI LOCAZIONE E SCALA}})$$

Teoria dei Valori Estremi

$$F_{a_n M_n + b_n}(x) = P(a_n M_n + b_n \le x) =$$
$$= P\left(M_n \le \frac{x - b_n}{a_n}\right) = P(M_n \le c_n x + d_n)$$

- la condizione $a_n M_n + b_n \to^{\mathrm{d}} L$ si può riscrivere come

$$F(c_n x + d_n)^n \to H(x) \text{ per ogni } x \in \mathbb{R}$$

- è possibile trovare le sequenze normalizzanti c_n, d_n quando F è dotata di densità f, usando

$$d_n = F^{-1}(1 - 1/n), \;\; c_n = \frac{1}{\lambda(d_n)}$$

dove $\lambda(x) = \frac{f(x)}{1-F(x)}$ (hazard rate)

- non è in generale possibile trovare sequenze normalizzanti per variabili aleatorie discrete ESEMPIO: BERNOULLI

ESEMPIO: DISTRIBUZIONE ESPONENZIALE

$F_i(x)$

$F(x) = 1 - e^{-\lambda x}, \quad x \geq 0$

$F^{-1}(x) = -\frac{1}{\lambda}\log(1-x) \qquad \lambda(x) = \lambda$

$d_n = F^{-1}\left(1-\frac{1}{n}\right) = -\frac{1}{\lambda}\log\left(1-1+\frac{1}{n}\right) = \frac{\log n}{\lambda}$

$c_n = \frac{1}{\lambda(d_n)} = \frac{1}{\lambda}$

... CONTINUA

$$F(c_n x + d_n)^n = \left(1 - e^{-\lambda(c_n x + d_n)}\right)^n$$

SE $c_n x + d_n > 0$

$$= \left(1 - e^{-\cancel{\lambda}\left(\frac{x}{\cancel{\lambda}} + \frac{\log n}{\cancel{\lambda}}\right)}\right)^n$$

$$= \left(1 + \frac{-e^{-x}}{n}\right)^n$$

$$\longrightarrow e^{-e^{-x}} \qquad n \to \infty$$

$c_n x + d_n > 0 \quad \forall x$, PER n GRANDE

$$\Leftrightarrow \frac{x}{\lambda} + \frac{\log n}{\lambda} > 0 \Leftrightarrow x > \underbrace{-\log n}_{\to -\infty}$$

ESEMPI

* $F_i(x) = F(x) = 1 - \left(\frac{\lambda}{\lambda + x}\right)^{\alpha}$ **PARETO**

  $x > 0$

  LIMITE : H^{II}_{α} FRECHET

* $F_i(x) = F(x) = x, \quad 0 < x < 1$ **UNIFORME (0,1)**

  LIMITE : H^{III}_{1} WEIBULL NEGATIVA

Teoria dei Valori Estremi

- Il dominio di attrazione di una distribuzione di valori estremi H è

$$\mathcal{D}_H = \{F \text{ funzione di ripartizione tale che esistono}$$

$$c_n,\, d_n \text{ per cui } F(c_n x + d_n)^n \to H(x) \text{ per ogni } x \in \mathbb{R}\}$$

- è possibile caratterizzare/dare delle proprietà comuni alle distribuzioni nei domini di attrazione in termini della coda destra della distribuzione, cioè del comportamento di $1 - F(x)$ quando $x \to \overline{x}$
- dominio di attrazione della Gumbel. $\mathcal{D}_{H^I}$ contiene distribuzioni a coda "leggera" ~~o senza coda~~: $1 - F(x)$ converge a 0 come un'esponenziale; tutti i momenti sono finiti; $P(\underset{\sim}{X} > x)$
  - ⋆ normale
  - ⋆ esponenziale e le sue generalizzazioni (gamma, Weibull)
  - ⋆ lognormale (anche se ha coda più pesante delle precedenti)
  - ⋆ ...

Teoria dei Valori Estremi

- dominio di attrazione della Frèchet. $\mathcal{D}_{H_\alpha^{II}}$ contiene distribuzioni a coda "pesante"; $1 - F(x)$ converge a 0 come una potenza: $F \in \mathcal{D}_{H_\alpha^{II}}$ se e solo se $\bar{x} = +\infty$ e

$$P(X_i > x) = 1 - F(x) = \frac{h(x)}{x^\alpha} = \frac{\text{"COSTANTE"}}{x^\alpha}$$

  dove h è a variazione lenta, $\lim_{x \to +\infty} \frac{h(tx)}{h(x)} = 1$ per ogni $t > 0$
  - ⋆ Pareto
  - ⋆ t di Student
  - ⋆ Cauchy
  - ⋆ ...

  Inoltre

$$E[\max(X, 0)^\beta] < +\infty \text{ se e solo se } \beta < \alpha$$

  Più piccolo è α (più grande è $\xi = 1/\alpha$), più pesante è la coda
- $\mathcal{D}_{H_\alpha^{III}}$ contiene distribuzioni senza coda in cui $\bar{x} < +\infty$
  - ⋆ uniforme
  - ⋆ beta
  - ⋆ ...

IL COMPORTAMENTO DI $1 - H_\alpha^{III}$ VICINO A $\bar{x}$ È SIMILE A QUELLO DI $1 - H_\alpha^{II}$ IN $+\infty$

FUNZIONI A VARIAZIONE LENTA

$$\lim_{x \to +\infty} \frac{h(t \cdot x)}{h(x)} = 1 \qquad \forall\, t > 0$$

$$1 - F(x) = P(X_i > x) \approx \frac{h(x)}{x^\alpha} \to 0$$

"LENTAMENTE"

ESEMPI

* $h(x) =$ COSTANTE $\Rightarrow h$ VAR. LENTA
* $\lim_{x \to +\infty} h(x) = \ell > 0$ //
* $h(x) = \log x$ //
* $h(x) = \log(\log x)$ //

- - - -

Teoria dei Valori Estremi

▷ stima: si estraggono i massimi dai blocchi

$$z_1, z_2, \ldots, z_m$$

con

$$\overbrace{\underbrace{x_1, \ldots x_n}_{\text{max: } z_1}}^{\text{blocco 1}}, \overbrace{\underbrace{x_{n+1}, \ldots x_{2n}}_{\text{max: } z_2}}^{\text{blocco 2}}, \ldots, \overbrace{\underbrace{x_{(j-1)n+1}, \ldots x_{jn}}_{\text{max: } z_j}}^{\text{blocco } j}, \ldots$$

▷ trade-off BIAS VS VARIANZA

n. di blocchi	n. di oss. in ogni blocco	approssimazione nel T. dei 3 Tipi (BIAS)	n. di oss. (VARIANZA) nel campione finale
↑	↓	peggiora	↑
↓	↑	migliora	↓

▷ Si modellizzano i massimi z_j con una GEV con cambio di locazione e scala, poi si stimano μ, σ, ξ NO RISCHIO DI MODELLO

Teoria dei Valori Estremi

- ▷ Approccio POT
  - ⋆ Siano $X_1, \dots, X_n, \dots$ i.i.d come prima; X è una v.a. con distribuzione F come gli X_i
  - ⋆ si considera, per una data soglia u, la v.a.

$$Y = (X - u | X > u)$$

SUPPORTO DI Y: $[0, \bar{x} - u)$

  - ⋆ riesce

$$= P(X - u \leq y | X > u) = P(u < X \leq u + y)$$

$$F_u(y) = \frac{\cancel{F(u) - F(u+y)}\; F(u+y) - F(u)}{1 - F(u)}$$

- ▷ Teorema (Balkema & de Haan (1974), Pickands (1975))
  se esistono $a_n > 0$, b_n tali che $a_n M_n + b_n \to^{\mathrm{d}} L$ con $L \sim G_\xi((\cdot - \mu)/\sigma)$, allora

  IPOTESI DEL THM DEI 3 TIPI

$$\text{per } u \to \bar{x}, \; F_u(y) \to W_\xi(y/\sigma_u) \text{ per ogni } 0 \leq y \leq \bar{x} - u,$$

  con W_ξ la distribuzione di Pareto Generalizzata (GPD) e $\sigma_u = \sigma + \xi(u - \mu)$

Teoria dei Valori Estremi

- la GPD è definita come:

$$W_\xi(y) = 1 - [1 + \xi y]^{-1/\xi}$$

con $y > 0$ and $1 + \xi y > 0$

- le tre distribuzioni contenute in questa espressione sono
  - $\xi \to 0$: Esponenziale
  - $\xi > 0$: Pareto
  - $\xi < 0$: Beta
- quindi eccessi oltre una soglia elevata si distribuiscono come una Pareto GENERALIZZATA

"U GRANDE"

$$P(X-u \leq y \mid X>u) \simeq 1-\left(1+\xi \frac{y}{\sigma_u}\right)^{-1/\xi}$$

Teoria dei Valori Estremi

▷ stima: il campione iniziale è

$$x_1, \ldots, x_n$$

- ⋆ si sceglie una soglia u
- ⋆ si estraggono le k osservazioni oltre u

NON È LA STATISTICA D'ORDINE!

$$x_{(1)}, \ldots, x_{(k)}$$

con $x_{(j)} > u$ per $j = 1, \ldots, k$

- ⋆ si costruiscono gli eccessi

$$y_1 = x_{(1)} - u, \ldots y_k = x_{(k)} - u$$

▷ dati $y_1, \ldots, y_k$, si usa il modello $y_j \sim W_\xi(\cdot/\sigma)$ e si stimano ξ, σ

▷ trade-off simile a quello del approccio block-maxima

POT : TRADE OFF (BIAS vs VARIANCE)

u	N° DI ECCESSI	APPROSSIMAZIONE DEL TMM (BIAS)	DIMENSIONE DEL CAMPIONE FINALE (VARIANZA)
↑	↓	↑	↓
↓	↑	↓	↑

Teoria dei Valori Estremi

- Proprietà della GPD: se $Y \sim W_\xi(\cdot/\sigma)$
  - $E[Y] = \frac{\sigma}{1-\xi}$ per $\xi < 1$
  - per ogni $u > 0$, $(Y - u|Y > u) \sim W_\xi(\cdot/(\sigma + \xi u))$
- Come scegliere la soglia?
  - le due proprietà precedenti implicano che se $Y \sim W_\xi(\cdot/\sigma)$ allora la mean excess function

$$e(u) = E[Y - u|Y > u] = \frac{\sigma + \xi u}{1 - \xi}, \qquad (\xi < 1)$$

funzione lineare della soglia!
  - si stima $e(u)$ con

$$\widehat{e}(u) = \frac{1}{k_u} \sum_{j=1}^{k_u} y_j = \frac{1}{k_u} \sum_{j=1}^{k_u} (x_{(j)} - u)$$

e si sceglie u tale che il grafico di $\widehat{e}$ diventa (approssimativamente) lineare da u in avanti

CALCOLO DI $E[Y]$

$$E[Y] = \int_0^{\bar{y}} P(Y > y)\,dy = \int_0^{\bar{y}} (1+\xi y)^{-1/\xi}\,dy$$

$$= \int_0^1 u \cdot u^{-\xi-1}\,du$$

$$= \int_0^1 u^{-\xi}\,du$$

$$= \left.\frac{u^{-\xi+1}}{-\xi+1}\right|_0^1$$

se $-\xi+1>0$
$\xi < 1$

$$= \frac{1}{1-\xi}$$

$(1+\xi y)^{-1/\xi} = u$

$y = \frac{u^{-\xi}-1}{\xi}$

$dy = -u^{-\xi-1}$

INVARIANZA DELLA PARETO (GEN) RISPETTO A ECCESSI CONDIZIONATI

$$Y \sim W_\xi$$

$$u > 0 \Rightarrow \boxed{(Y-u \mid Y>u) \sim W_\xi\left(\frac{\cdot}{\sigma+\xi u}\right)}$$

$$P(Y-u \le y \mid Y>u) = \frac{P(u<Y\le u+y)}{P(Y>u)}$$

$$= \frac{\left(1+\xi\frac{u}{\sigma}\right)^{-1/\xi} - \left(1+\xi\frac{u+y}{\sigma}\right)^{-1/\xi}}{\left(1+\xi\frac{u}{\sigma}\right)^{-1/\xi}}$$

... CONTINUA

$$= 1 - \left(\frac{1 + \xi \frac{u+y}{\sigma}}{1 + \xi \frac{u}{\sigma}} \right)^{-1/\xi}$$

$$= 1 - \left(\frac{\sigma + \xi(u+y)}{\sigma + \xi u} \right)^{-1/\xi}$$

PARETO $\longrightarrow$ GENERALIZZATA

$$= 1 - \left(1 + \xi \frac{y}{\sigma + \xi u} \right)^{-1/\xi}$$

STESSO PARAMETRO DI FORMA

CAMBIA IL PARAMETRO DI SCALA

Teoria dei Valori Estremi

- Calcolo di VaR e ES con l'approccio POT ($X = L$)
- VaR:
  - la distribuzione di X è approssimativamente, per $x > u$,

$$F(x) = 1 - P(X > x|X > u)P(X > u) \approx 1 \cancel{- \frac{k}{n} W_\xi\left(\frac{x-u}{\sigma}\right)},$$

$$1 - \frac{k}{n}\left(1 - W_\xi\left(\frac{x-u}{\sigma}\right)\right)$$

si calcola il VaR risolvendo $F(x) = \alpha$

- ES:
  - usando la mean excess function,

$$\mathrm{ES}_\alpha = E[X|X > \mathrm{VaR}_\alpha] = \mathrm{VaR}_\alpha + E[X - \mathrm{VaR}_\alpha\,|X > \mathrm{VaR}_\alpha]$$

$$= \mathrm{VaR}_\alpha + e(\mathrm{VaR}_\alpha - u) = \mathrm{VaR}_\alpha + \frac{\sigma + \xi(\mathrm{VaR}_\alpha - u)}{1 - \xi}$$

APPROSSIMAZIONE PER $F(x)$, $x > u$

$$F(x) = P(X \le x) = 1 - P(X > x)$$

$$= 1 - P(X > x \mid X > u)P(X > u)$$

$$- \underbrace{P(X > x \mid X \le u)}_{=0\ ,\ u < x} P(X \le u)$$

$$= 1 - \underbrace{P(X > u)}_{\approx \frac{k}{n}} \left(1 - \underbrace{P(X \le x \mid X > u)}_{P(X - u \le x - u \mid X > u)}\right)$$

$$P(X - u \le x - u \mid X > u) \approx W_\xi\left(\frac{x-u}{\sigma}\right) = 1 - \left(1 + \xi \frac{x-u}{\sigma}\right)^{-1/\xi}$$

k = N. DI ECCESSI

n = N. DI OSSERVAZIONI

CALCOLO DEL VaR $(X = L)$

$$F(x) = \alpha$$

$$1 - \frac{k}{n}\left(1 + \xi\,\frac{x - u}{\sigma}\right)^{-1/\xi} = \alpha$$

$$\boxed{x = \mathrm{VaR}_\alpha(L) = u + \frac{\sigma}{\xi}\left(\left(\frac{k}{n}(1-\alpha)\right)^{-\xi} - 1\right)}$$

CALCOLO DI ES $(X=L)$

~~Teoria dei Valori Estremi~~ $VaR_\alpha(X) > u$

$$Y = (X - u \mid X > u)$$

$$Y' = (X - VaR_\alpha(X) \mid X > VaR_\alpha(X))$$

OSS: $Y' = (Y - (VaR - u) \mid Y > VaR - u)$

$$ES_\alpha(X) = E(X \mid X > VaR_\alpha(X)) =$$

$$= VaR_\alpha(X) + E(\overbrace{X - VaR_\alpha(X) \mid X > VaR_\alpha(X)}^{Y'})$$

$$= VaR_\alpha(X) + \frac{\sigma + \xi(VaR_\alpha(X) - u)}{1 - \xi}$$