

## VALUE-AT-RISK: HISTORICAL SIMULATION

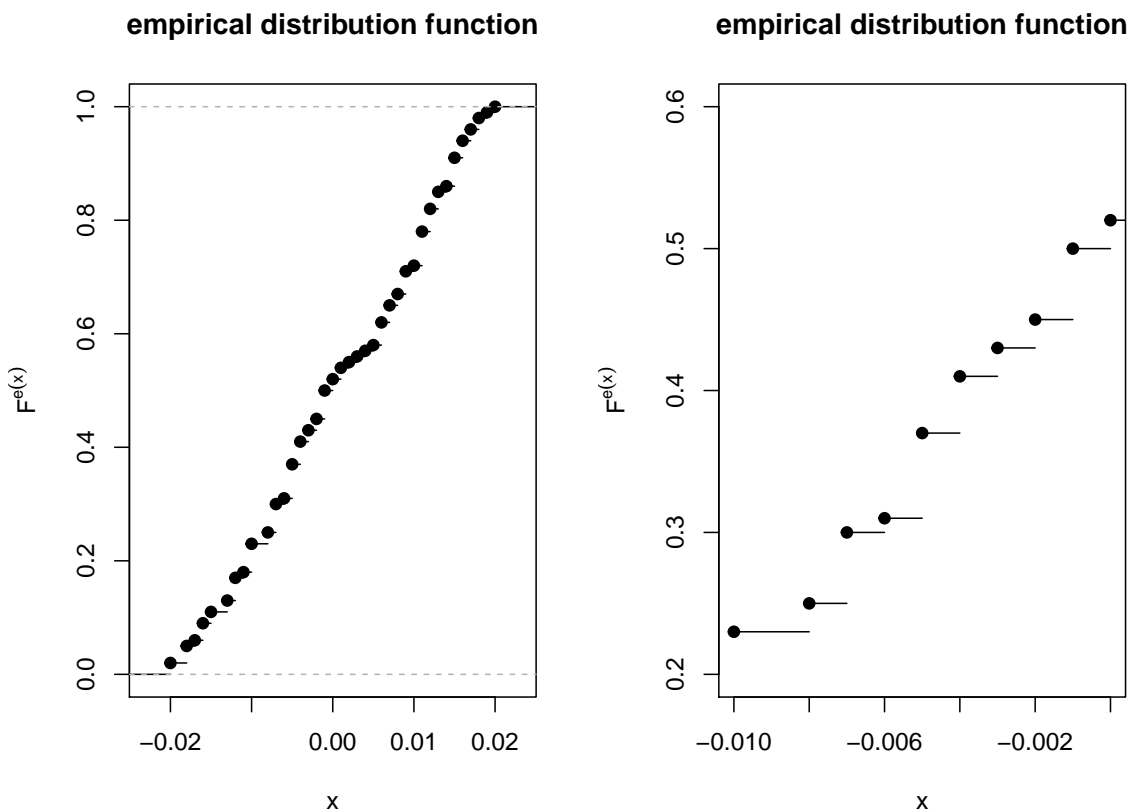
- ▷ approccio non parametrico: **historical simulation**
  - ★ si usa la **funzione di ripartizione empirica** costruita da un campione casuale  $l_1, \dots, l_m$  osservato da  $L$

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

- ★ indicato con  $l_{(1)} \leq l_{(2)} \leq \dots \leq l_{(m)}$  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  $x_{(i)}$
- ★ se gli  $l_i$  non sono tutti distinti: se  $l_{(i-1)} < l_{(i)} = l_{(i+1)} = \dots = l_{(i+k-1)} < l_{(i+k)}$  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

310

## VALUE-AT-RISK: HISTORICAL SIMULATION



311

## VALUE-AT-RISK: HISTORICAL SIMULATION

- ▷ 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

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

- ★ se  $\alpha m$  è un intero, allora

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

- ★ in generale, (necessariamente) un intero, allora

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

dove  $\lceil x \rceil =$  minimo intero maggiore o uguale a  $x$

312

## VALUE-AT-RISK: HISTORICAL SIMULATION

- ▷ 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  $x$

- ★ si pone

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

dove  $0 \leq \gamma \leq 1$  è un peso

- ★ classica scelta:  $\gamma = \alpha m - \lfloor \alpha m \rfloor \rightsquigarrow$  quantile continuo in  $\alpha$   
 ★ per  $\gamma = 0$  se  $\alpha m$  intero,  $\gamma = 1$  se  $\alpha m$  non intero, si ritorna alla definizione precedente

- ▷ Per l'expected shortfall,

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

dove il secondo termine additivo dentro parentesi viene a volte trascurato

313

## EXPECTED SHORTFALL: 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  $\text{VaR}_{0.9}(L) = l_{(9)} = 3\%$

★ per  $\beta > 0.9$ , riesce  $\text{VaR}_\beta = 4\%$

★ l'expected shortfall è

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

★ si noti che  $E[L|L \geq \text{VaR}_{0.9}] = 3.5\%$

314

## MISURE DI RISCHIO

- ▷ approccio **assiomatico** alle misure di rischio: stabilire proprietà ragionevoli che una misura di rischio  $\rho : \mathcal{L} \rightarrow \mathbb{R}$  potrebbe soddisfare

★ **no rip-off**:

$\rho(L) \leq$  estremo superiore di  $L = \inf\{x \in \mathbb{R} : F_L(x) = 1\}$  per ogni  $L \in \mathcal{L}$ ;

non si dovrebbe allocare più capitale della perdita massima

★ **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  $L_1$

315

## 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  $L \in \mathcal{L}$ ,  $c \in \mathbb{R}$ ;  
aggiungendo una perdita certa  $c$  ad  $L$ , si deve aumentare il capitale dello stesso ammontare
  - ★ **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**
  - ★ **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

316

## MISURE DI RISCHIO

- ▷ una misura di rischio che soddisfa la **monotonia, invarianza per traslazioni, positiva omogeneità e sub-additività** si dice **coerente**
- ▷ una misura di rischio che soddisfa la **monotonia, invarianza per traslazioni e convessità** si dice **convessa**
- ▷ ogni misura di rischio positivamente omogenea è subadditiva se e solo se verifica la convessità  $\rightsquigarrow$  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  $\rightsquigarrow$  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
- ▷ le misure di rischio basate su scenari generalizzati (con pesi tutti unitari) sono coerenti

317

## SUBADDITIVITÀ DELL'EXPECTED SHORTFALL

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

- ★  $I_i = 1_{\{L_i \geq \text{VaR}_\alpha(L_i)\}}$ ,  $i = 1, 2$ ,  $I_{12} = 1_{\{L \geq \text{VaR}_\alpha(L)\}}$
- ★ si osservi che

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

- ★ riesce

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

(considerare separatamente i casi  $L_i \geq \text{VaR}_\alpha(L_i)$  e  $L_i < \text{VaR}_\alpha(L_i)$ )

- ★ concludere che

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

318

## 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

$\Gamma : [0, 1] \rightarrow [0, 1]$  ( $\Gamma$  non decrescente, continua a dx,  $\Gamma(0) = 0$ ,  $\Gamma(1) = 1$ ), si pone

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

- ★ idea: pesare i capitali a ogni livello di confidenza con la misura  $\Gamma$
- ★  $\text{VaR}_\alpha$ :  $\Gamma$  concentrata in  $\alpha$

▷ quando  $\Gamma$  è assolutamente continua con densità  $\gamma$ ,

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

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

319

## MISURE DI RISCHIO BASATE SU PERDITE

- ▷ Una misura di rischio convessa non (positivamente omogenea)  $\rightsquigarrow$  non coerente
- ▷  $\ell : \mathbb{R} \rightarrow \mathbb{R}$  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

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

- ▷  $\rho$  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  $z^\lambda$ )

320

## VALUE-AT-RISK E SUBADDITIVITÀ

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

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

per qualche  $\mu \in \mathbb{R}$ ,  $\sigma^2 \geq 0$ , per ogni  $a_1, \dots, a_d \in \mathbb{R}$  (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

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

321

## 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:
- ★ generalizzano al caso multivariato le proprietà di **simmetria** della normale
  - ★ distribuzioni **marginali** di distribuzioni ellittiche sono ancora ellittiche
  - ★ trasformazioni **affini** di distribuzioni ellittiche sono ancora ellittiche
  - ★ distribuzioni **condizionate** di distribuzioni ellittiche sono ancora ellittiche

322

## VALUE-AT-RISK E SUBADDITIVITÀ

- ▷ un vettore  $\mathbf{X} = [X_1, \dots, X_d]^\top$  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

$$f(\mathbf{x}) = \tilde{f}(x_1^2 + \dots + x_d^2), \text{ per ogni } \mathbf{x} \in \mathbb{R}^d$$

dove  $\tilde{f}: [0, +\infty) \rightarrow [0, \infty)$ , cioè la densità è costante sulle sfere.

- ★  $t_\nu$  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

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

323

## VALUE-AT-RISK E SUBADDITIVITÀ

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

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

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  $\alpha \geq 0.5$

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

★ **Value-at-Risk è consistente con la varianza**

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

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

324

## 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)

325



## PROFITTO/PERDITA

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

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

variazione di valore dell'attività

- ▷  $P\&L_{t,T}$  rappresenta il guadagno/perdita se si prende una posizione lunga (acquista) il sottostante in  $t$  e si liquida la posizione in  $T$
- ▷ interpretazione

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

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

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

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

326

## PROFITTO/PERDITA

- ▷ il **rendimento** viene usualmente misurato su base
- ★ semplice o composto
  - ★ periodale o annuo
- ▷ **Rendimento periodale semplice**

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

cioè

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

- ★ su base annua, il rendimento equivalente è

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

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

327

## PROFITTO/PERDITA

- ▷ Rendimento periodale composto

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

cioè

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

- ★ su base annua, il rendimento equivalente è

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

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

328

## PROFITTO/PERDITA

- ▷ Relazioni tra rendimenti semplici e composti

$$I = e^R - 1$$

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

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

$$R \approx I$$

e viceversa (formula di Taylor)

- ▷ 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)

329

## VAR PER PERDITE E PER RENDIMENTI

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

$$\text{VaR}_\alpha(L) = \inf\{x \in \mathbb{R} : \text{Prob}[-\text{P\&L} \leq x] \geq \alpha\} = F_{-\text{P\&L}}^{-1}(\alpha)$$

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

$$\text{VaR}_\alpha(-R) = \inf\{x \in \mathbb{R} : \text{Prob}[-R \leq x] \geq \alpha\} = F_{-R}^{-1}(\alpha)$$

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

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

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

- ▷ simili relazioni se si utilizzano i rendimenti semplici

330

## 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  $(t_0, t_1), \dots, (t_{n-1}, t_n)$ , con  $t_0 = t < t_1 < \dots < t_{n-1} < t_n = T$

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

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

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

- ★ nel caso di rendimenti composti/annui?

331

## 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
- ▷ 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

332

## TASSI DI CAMBIO

- ▷ vantaggio dei rendimenti composti: sia  $V(t)$  il tasso di cambio f/d (foreign/domestic)
- ▷  $V(t) =$  quantità 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:

$$R_{t,T}^d = \log \frac{V(T)}{V(t)}$$

- ▷ rendimento composto per un investitore straniero:

$$R_{t,T}^f = \log \frac{V(t)}{V(T)} = -R_{t,T}^d$$

333

## RENDIMENTI COMPOSTI: PROPRIETÀ TEMPORALI

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

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

- ▷ segue che

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

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

- ▷ Se i rendimenti sono correlati?

334

## RENDIMENTI COMPOSTI: PROPRIETÀ TEMPORALI

- ▷ nel caso di intervalli di **ugual ampiezza** e **distribuzioni stazionarie**:

- ★  $t_i - t_{i-1} = \Delta$  per  $i = 1, \dots, n$ , 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”**:

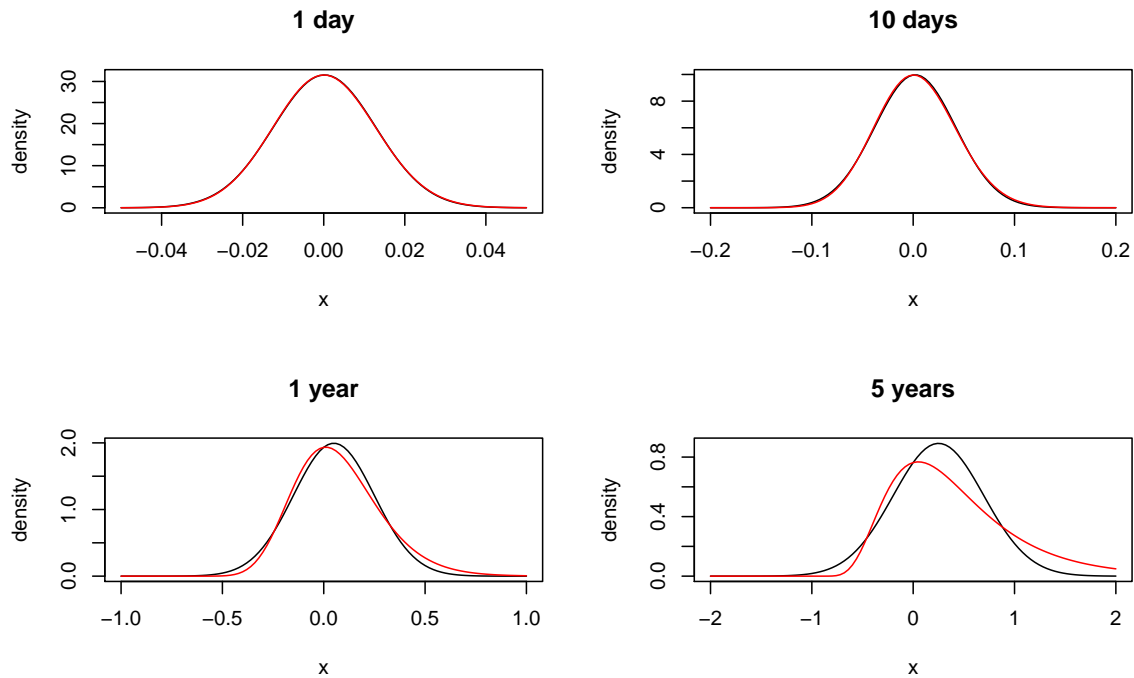
$$\text{sd}[R_{t,t+n\Delta}] = \sigma\sqrt{n}$$

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)

335

## RENDIMENTI COMPOSTI: PROPRIETÀ TEMPORALI



336

## P&L DI PORTAFOGLIO

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

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

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

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

- ★ rendimento semplice e composto di portafoglio

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

337

## P&L DI PORTAFOGLIO

▷ percentuale di ricchezza investita nel titolo  $j$ :

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

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

- ★ rendimento semplice di portafoglio è la media aritmetica ponderata dei rendimenti semplici (“**additivity across assets**”)

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

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

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

338

## P&L DI PORTAFOGLIO

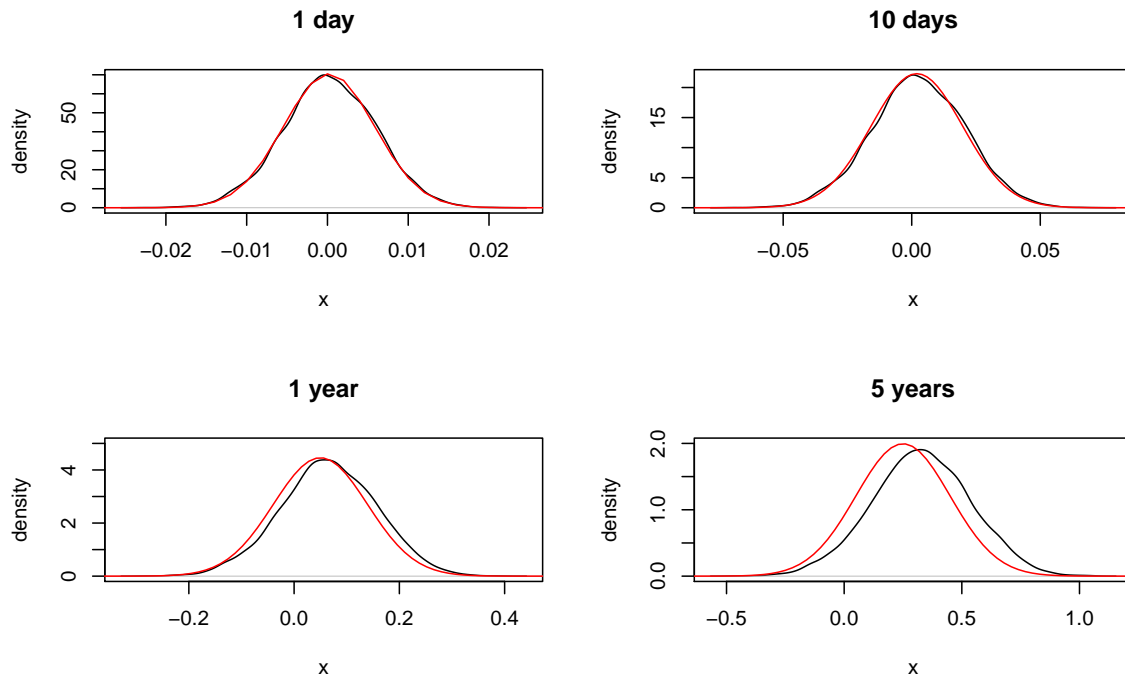
▷ Se i rendimenti delle attività componenti il portafoglio sono “piccole”, allora vale l’approssimazione (usare  $e^x \approx 1 + x$  e  $\log(1 + x) \approx x$  quando  $x \rightarrow 0$ )

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

- ★ se la distribuzione congiunta di  $R_{t,T}^1, \dots, R_{t,T}^N$  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

339

## RENDIMENTI DI PORTAFOGLIO



340

## 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
- ★  $E[R_{t, t+n\Delta}] = n\mu$ ,  $\text{sd}[R_{t, t+n\Delta}] = \sigma\sqrt{n}$
  - ★ il Value-at-Risk sull'orizzonte temporale  $(t, t + n\Delta)$  è

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

Se  $\mu = 0$ , allora

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

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

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

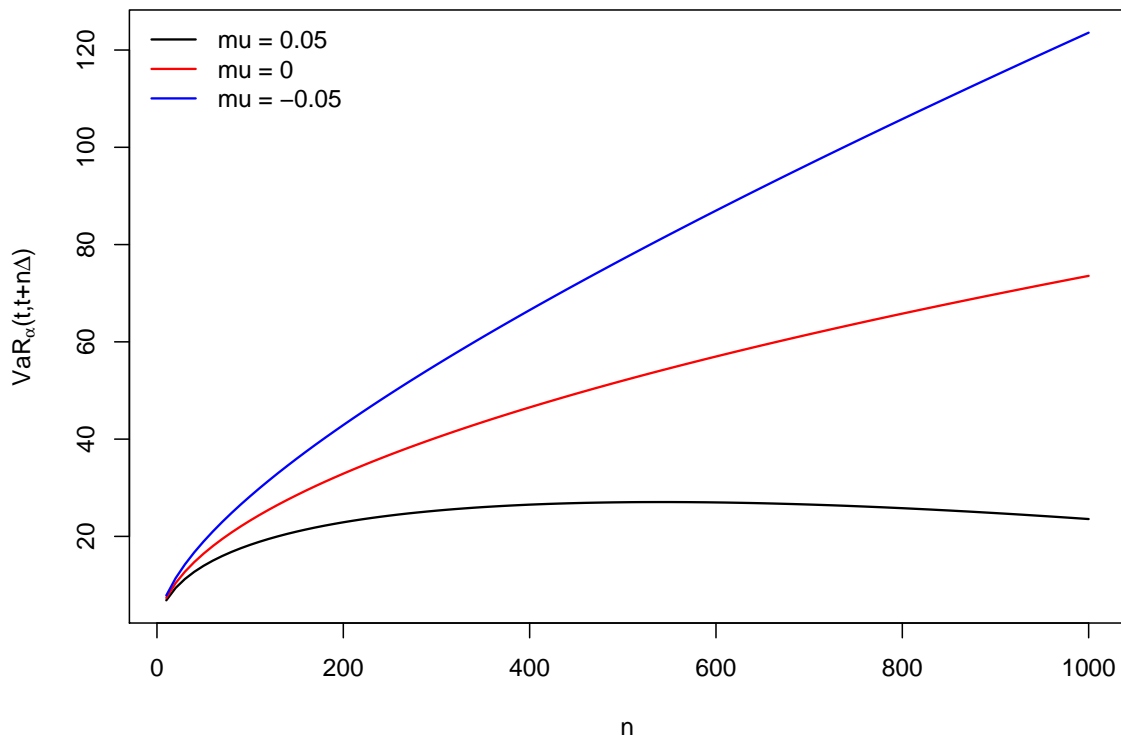
$$\text{VaR} \uparrow n \text{ se } \mu \leq 0$$

VaR prima crescente, poi decrescente, se  $\mu > 0$

341



## VALUE-AT-RISK



342

## VALUE-AT-RISK: APPROCCIO PARAMETRICO

### ▷ 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?
- \* rischio di parametro: il metodo richiede la stima dei parametri  $\mu$ ,  $\sigma$  (e  $\nu$  nel caso  $t$  di Student)  $\Rightarrow$  intervalli di confidenza per il Value-at-Risk
- \* rischio di modello: il metodo richiede la scelta di un modello  $\Rightarrow$  ogni modello è sbagliato

343

## 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**

344

## 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:

$$l_{(1)} = -2\%, l_{(2)} = l_{(3)} = -1\%, l_{(4)} = 0\%, l_{(5)} = 0.5\%,$$

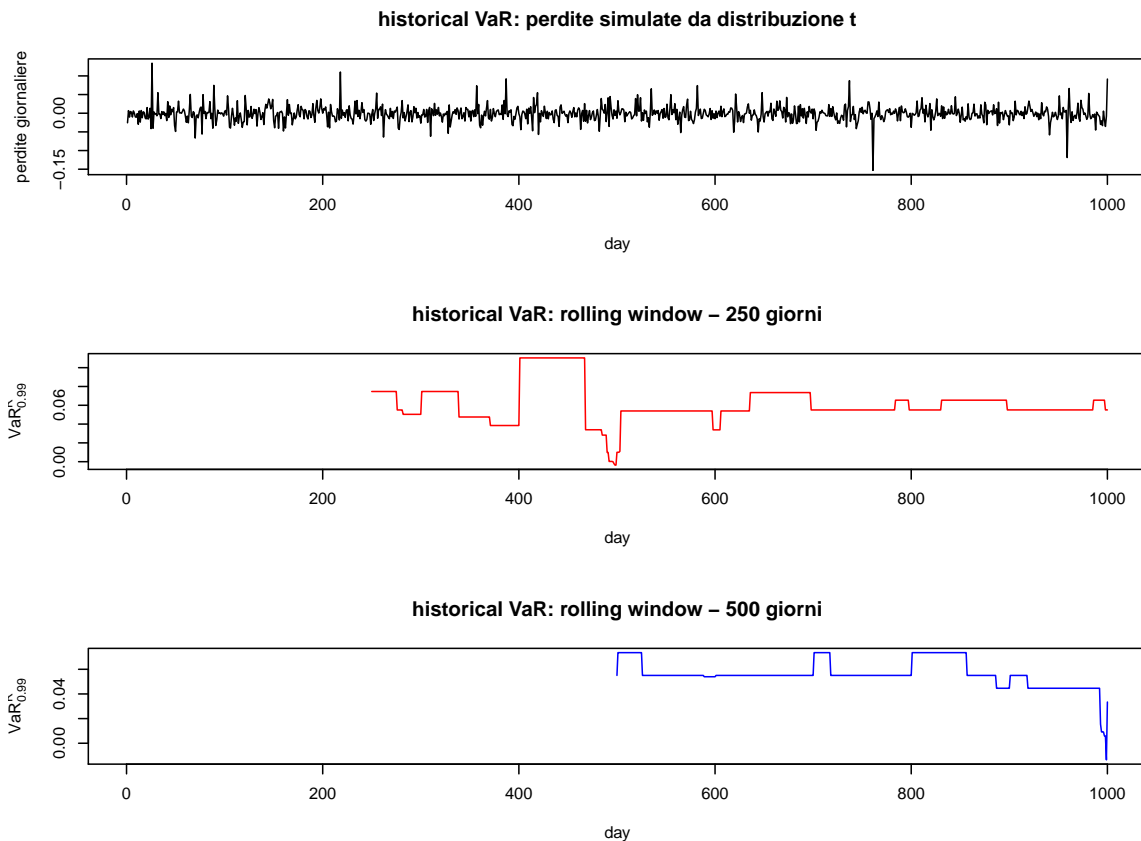
$$l_{(6)} = l_{(7)} = 1\%, l_{(8)} = 2.5\%, l_{(9)} = 3\%, l_{(10)} = 4\%$$

si trova  $\text{VaR}_{0.9} = l_{(9)} = 3\%$

- ★ 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  $l_{11} \leq 3\%$  allora  $\text{VaR}_{0.9} = 3\%$ ; se  $3\% < l_{11} \leq 4\%$  allora  $\text{VaR}_{0.9} = l_{11}$ ; se  $l_{11} > 4\%$  allora  $\text{VaR}_{0.9} = 4\%$

345

## VALUE-AT-RISK: HISTORICAL SIMULATION



346

## WEIGHTED HISTORICAL VALUE-AT-RISK (WHVaR)

- ▷ 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  $l_1, \dots, l_m$ , il peso di  $l_i$  è

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

con  $0 < \eta < 1$  (tipicamente  $\eta > 0.9$ )

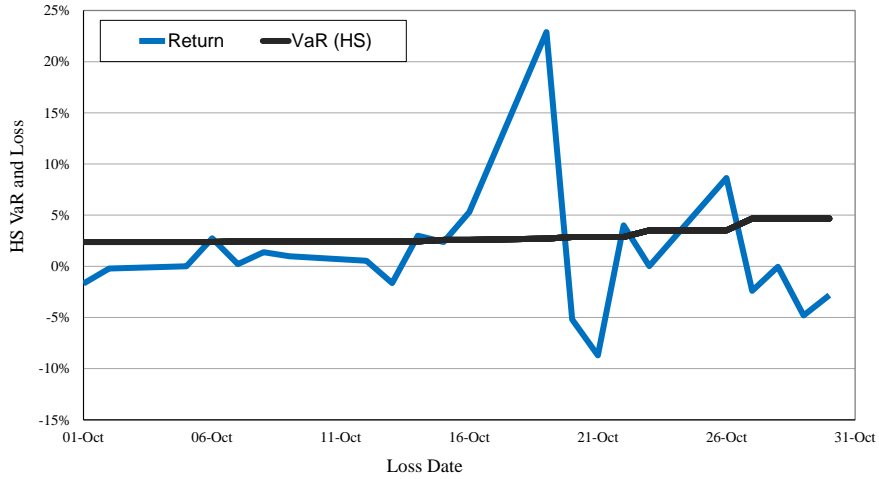
- ★ in questo modo l'osservazione più vecchia che esce dal campione ha poco peso, mentre ha peso massimo la più recente
- ★ il ruolo di  $m$  è meno importante
- ★ si calcola poi il percentile della distribuzione (tenendo conto dei pesi)

347

# WEIGHTED HISTORICAL VALUE-AT-RISK

15

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



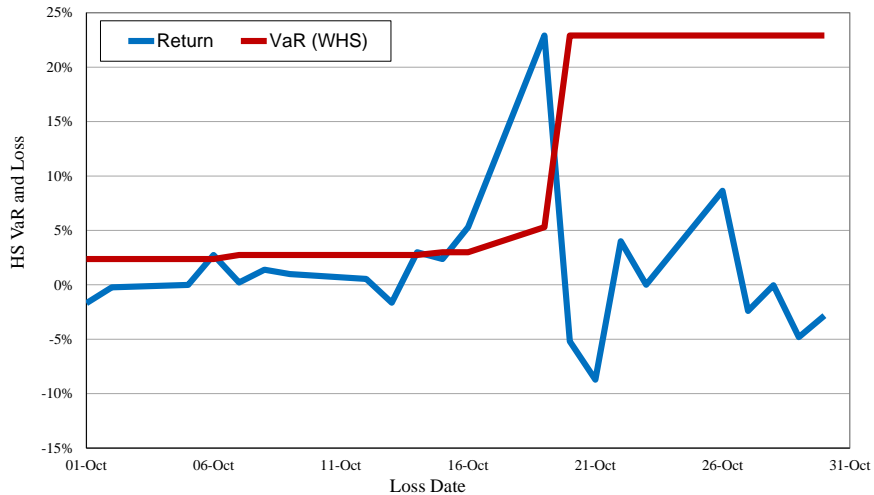
Elements of Financial Risk Management Second Edition © 2012 by Peter Christoffersen

348

# WEIGHTED HISTORICAL VALUE-AT-RISK

16

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



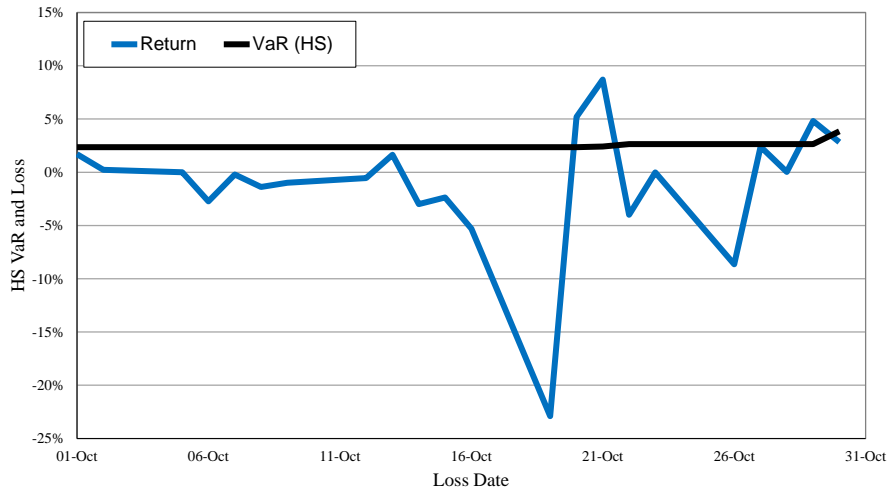
Elements of Financial Risk Management Second Edition © 2012 by Peter Christoffersen

349

# WEIGHTED HISTORICAL VALUE-AT-RISK

17

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



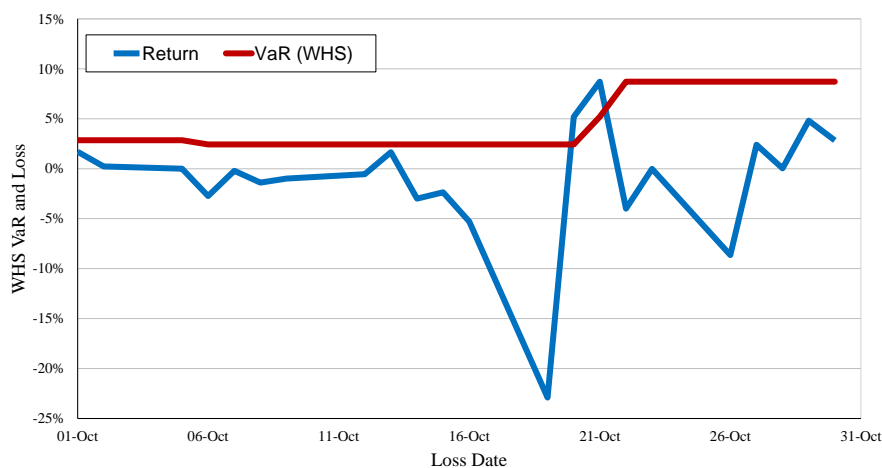
Elements of Financial Risk Management Second Edition © 2012 by Peter Christoffersen

350

# WEIGHTED HISTORICAL VALUE-AT-RISK

18

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

351

## 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
  - ★ idea: (ri)estrarre **con reinserimento** un campione da quello dato, di uguale lunghezza
  - ★ calcolare la statistica di interesse sul nuovo campione
  - ★ ripetere la procedura un numero elevato di volte
  - ★ si ottiene così la distribuzione campionaria della statistica di interesse  $\Rightarrow$  media, intervalli di confidenza, ...

352

## VALUE-AT-RISK: BOOTSTRAP

- ▷ **bootstrap VaR, 1 periodo**
  1. è dato un campione  $r_1, \dots, r_m$  di rendimenti uniperiodali (e.g. giornalieri)
  2. si costruisce un nuovo campione di medesima lunghezza  $m$  via **estrazione con rimpiazzo**, indicato con  $\hat{r}_{1,1}, \dots, \hat{r}_{1,m}$
  3. si calcola (via **simulazione storica**) il Value-at-Risk dal nuovo campione, indicato con  $\widehat{\text{VaR}}_{\alpha,1}(-R)$
  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  $\widehat{\text{VaR}}_{\alpha,1}(-R), \dots, \widehat{\text{VaR}}_{\alpha,M}(-R)$
  6. si calcola poi la **stima di bootstrap** del VaR via

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

353

## 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  $\widehat{\text{VaR}}_{\alpha, \text{bootstrap}}(-R)$
  - ★ si calcolano i Value-at-Risk per ogni campione bootstrap come in 1.-5.,  $\widehat{\text{VaR}}_{\alpha, 1}(-R), \dots, \widehat{\text{VaR}}_{\alpha, M}(-R)$
  - ★ si sceglie un livello di confidenza  $\alpha'$ , non collegato con il livello di confidenza del Value-at-Risk  $\alpha$ !
  - ★ si calcolano poi il  $(1 - \alpha')/2$  e il  $(1 + \alpha')/2$  quantili della distribuzione  $\widehat{\text{VaR}}_{\alpha, 1}(-R), \dots, \widehat{\text{VaR}}_{\alpha, M}(-R)$
- ▷ il metodo funziona bene se il campione iniziale non è **troppo piccolo** e se proviene da una distribuzione **ragionevolmente simmetrica**

354

## 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
  - ★ 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 bootstrap 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

355

## VALUE-AT-RISK: BOOTSTRAP

- ▷ **bootstrap VaR,  $n$  periodi:**  $(t, T) = (t, t + n\Delta)$
1. è dato un campione  $r_1, \dots, r_m$  di rendimenti uniperiodali (e.g. giornalieri)
  2. si costruisce un nuovo campione di lunghezza  $n$  via **estrazione con rimpiazzo**, indicato con  $\widehat{r}_{1,1}(1), \dots, \widehat{r}_{1,1}(n)$
  3. si ottiene un rendimento sul periodo  $(t, t + n\Delta)$  via

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

4. si ripetono 2.-3.  $m$  volte in maniera da ottenere un campione  $\widehat{r}_{1,1}(t, t + n\Delta), \dots, \widehat{r}_{1,m}(t, t + n\Delta)$  di rendimenti  $n$ -periodali
5. si calcola (via **simulazione storica**) il Value-at-Risk dal campione multi periodale, indicato con  $\widehat{\text{VaR}}_{\alpha,1}(-R)$
6. si ripetono 2.-4. un numero scelto  $M$  di volte, ottenendo  $M$  stime di Value-at-Risk, indicati con  $\widehat{\text{VaR}}_{\alpha,1}(-R), \dots, \widehat{\text{VaR}}_{\alpha,M}(-R)$
7. si calcola poi la **stima di bootstrap** del VaR via

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

356

## VALUE-AT-RISK: BOOTSTRAP

- ▷ **bootstrap VaR,  $n$  periodi:**  $(t, T) = (t, t + n\Delta)$
- ★ similmente si aggiungono intervalli di confidenza
  - ★ l'ipotesi fondamentale è che i rendimenti siano indipendenti su periodi diversi  $\Rightarrow$  estrazioni indipendenti dallo stesso campione
  - ★ 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)

357



## 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
- ★ 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

358

## 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*

359

## 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.”*

360

## TEORIA DEI VALORI ESTREMI

- ▷ formalizzando i concetti visti prima, si consideri una sequenza di variabili aleatorie

$$X_1, X_2, \dots, X_n, \dots$$

**indipendenti 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)
- ▷ 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

361

## TEORIA DEI VALORI ESTREMI

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

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

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

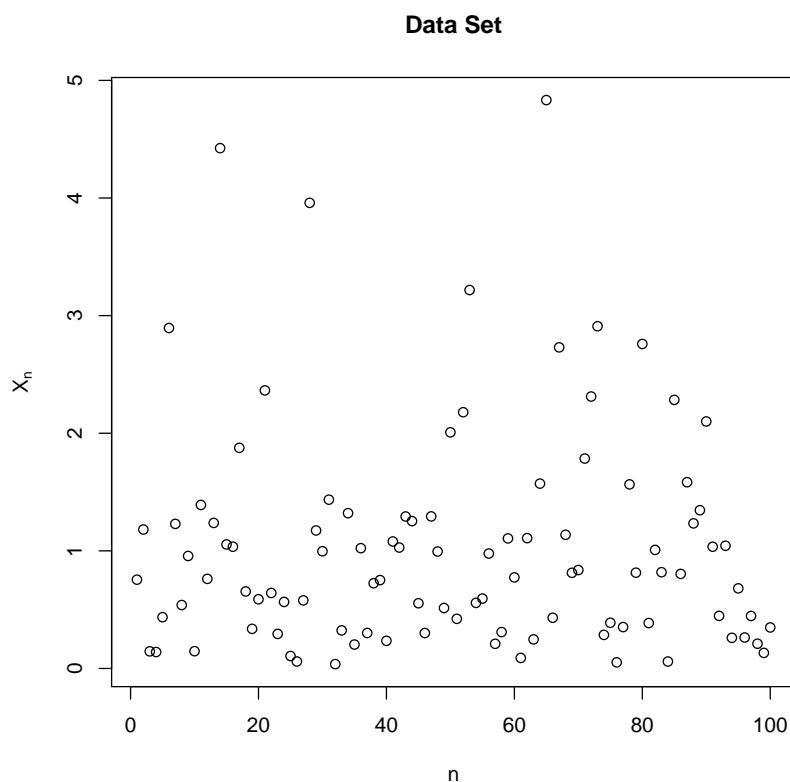
- ▷ si possono considerare **minimi**, dal momento che

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

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

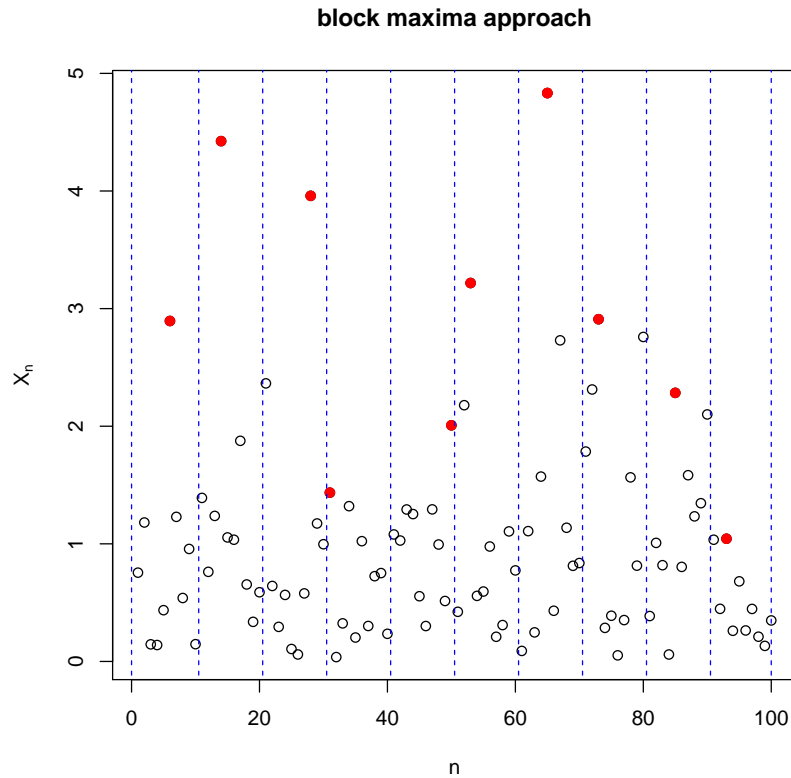
362

## TEORIA DEI VALORI ESTREMI



363

## TEORIA DEI VALORI ESTREMI



364

## 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

$$* P[X_n > u + y | X_n > u], y \geq 0$$

$$* E[X_n - u | X_n > u]$$

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

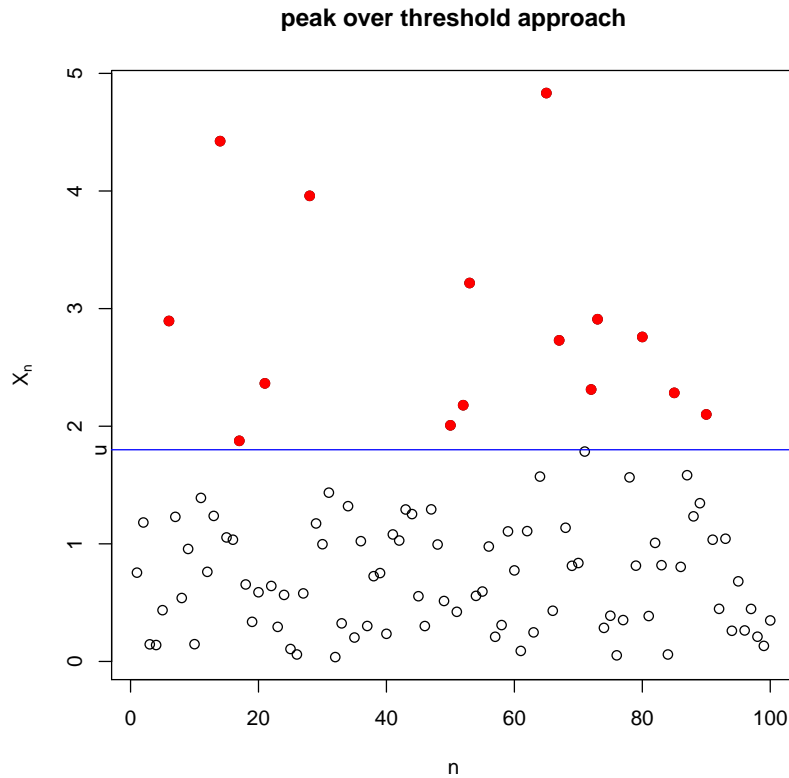
\* la scelta di  $u$  è critica

\* si possono considerare eccessi sotto una soglia  $u$

$$P[X_n \leq u - y | X_n \leq u], y \geq 0$$

365

# TEORIA DEI VALORI ESTREMI



366

# 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  $\rightsquigarrow$  l'**errore di modello** è meno importante
- ★ l'EVT fornisce una chiara indicazione sulla natura della coda della distribuzione dei dati

367

## TEORIA DEI VALORI ESTREMI

### ▷ difetti della EVT

- ★ teoria asintotica  $\rightsquigarrow$  richiede larghezza campionaria sufficientemente grande perchè l'approssimazione sia efficace  $\rightsquigarrow$  **trade-off** tra
  - bias (distorsione del modello)
  - varianza (precisione degli stimatori)
- ★ EVT di base assume osservazioni iid  $\rightsquigarrow$  spesso non soddisfatta

### ▷ estensioni della EVT

- ★ EVT può essere estesa a osservazioni **stazionarie** (dipendenti)
  - ★ tuttavia, i dati spesso esibiscono **trend** e/o **stagionalità**  $\rightsquigarrow$  non stazionarie
  - ★ studio congiunto di due o più fenomeni correlati, eg
    - livello del mare (velocità del vento, quantità di piogge, neve ...)
    - in due località diverse
    - velocità del vento e quantità di pioggia in una data località
    - serie finanziarie diverse — FX e azionario
- $\rightsquigarrow$  EVT **multivariata**

368

## 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**

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

- ▷ si può estendere la considerazione alle  $k$  osservazioni più grandi

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

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

369

## TEORIA DEI VALORI ESTREMI

- ▷ l'ipotesi di base è che le variabili aleatorie  $X_1, \dots, X_n, \dots$  siano **indipendenti 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  $n \rightarrow +\infty$  è facilmente descritto:  $M_n$  **converge in distribuzione** alla variabile aleatoria degenera in  $\bar{x} := \inf\{x \in \mathbb{R} : F(x) = 1\}$ :

$$M_n \rightarrow^d \bar{x}$$

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

370

## TEORIA DEI VALORI ESTREMI

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

$$a_n M_n + b_n \rightarrow^d L.$$

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**

371

## TEORIA DEI VALORI ESTREMI

▷ le distribuzioni dei valori estremi sono

★ Tipo I - Gumbel

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

★ Tipo II - Fréchet

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

con  $\alpha > 0$

★ Tipo III - Weibull negativa

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

con  $\alpha > 0$

372

## 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)**

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

★  $\xi \rightarrow 0$  si trova la Gumbel

★  $\xi = 1/\alpha > 0$  si trova la Fréchet

★  $\xi = -1/\alpha < 0$  si trova la Weibull negativa

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

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

373



## TEORIA DEI VALORI ESTREMI

- ▷ la condizione  $a_n M_n + b_n \xrightarrow{d} L$  si può riscrivere come

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

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

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

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

- ▷ non è in generale possibile trovare sequenze normalizzanti per variabili aleatorie discrete

374

## TEORIA DEI VALORI ESTREMI

- ▷ Il **dominio di attrazione** di una distribuzione di valori estremi  $H$  è

$$\mathcal{D}_H = \{F \text{ funzione di ripartizione tale che esistono } c_n, d_n \text{ per cui } F(c_n x + d_n)^n \rightarrow H(x) \text{ per ogni } x \in \mathbb{R}\}$$

- ▷ 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  $x \rightarrow \bar{x}$
- ▷ **dominio di attrazione della Gumbel**.  $\mathcal{D}_{HI}$  contiene distribuzioni a **coda "leggera" o senza coda**:  $1 - F(x)$  converge a 0 come un'esponenziale; tutti i momenti sono finiti;
- ★ normale
  - ★ esponenziale e le sue generalizzazioni (gamma, Weibull)
  - ★ lognormale (anche se ha coda più pesante delle precedenti)
  - ★ ...

375

## TEORIA DEI VALORI ESTREMI

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

$$1 - F(x) = \frac{h(x)}{x^\alpha}$$

dove  $h$  è a **variazione lenta**,  $\lim_{x \rightarrow +\infty} \frac{h(tx)}{h(x)} = 1$  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  $\alpha$  (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
  - ★ ...

376

## TEORIA DEI VALORI ESTREMI

- ▷ **stima:** si estraggono i massimi dai blocchi

$$z_1, z_2, \dots, z_m$$

con

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

- ▷ **trade-off**

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

- ▷ Si modellizzano i massimi  $z_j$  con una GEV con cambio di locazione e scala, poi si stimano  $\mu, \sigma, \xi$

377

## 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.

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

- ★ riesce

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

### ▷ Teorema (Balkema & de Haan (1974), Pickands (1975))

se esistono  $a_n > 0$ ,  $b_n$  tali che  $a_n M_n + b_n \xrightarrow{d} L$  con  $L \sim G_\xi((\cdot - \mu)/\sigma)$ , allora

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

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

378

## 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

- ★  $\xi \rightarrow 0$ : Esponenziale
- ★  $\xi > 0$ : Pareto
- ★  $\xi < 0$ : Beta

### ▷ quindi **eccessi oltre una soglia elevata si distribuiscono come una Pareto**

379

## TEORIA DEI VALORI ESTREMI

- ▷ **stima:** il campione iniziale è

$$x_1, \dots, x_n$$

- ★ si **sceglie una soglia  $u$**
- ★ si **estraggono le  $k$  osservazioni oltre  $u$**

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

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

- ★ si **costruiscono gli eccessi**

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

- ▷ dati  $y_1, \dots, y_k$ , si usa il modello  $y_j \sim W_\xi(\cdot/\sigma)$  e si stimano  $\xi, \sigma$
- ▷ trade-off simile a quello del approccio block-maxima

380

## 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**

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

**funzione lineare della soglia!**

- ★ si stima  $e(u)$  con

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

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

381

## TEORIA DEI VALORI ESTREMI

▷ Calcolo di VaR e ES con l'approccio POT ( $X = L$ )

▷ **VaR:**

★ la distribuzione di  $X$  è approssimativamente, per  $x > u$ ,

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

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

▷ **ES:**

★ usando la mean excess function,

$$\begin{aligned} \text{ES}_\alpha &= E[X | X > \text{VaR}_\alpha] = \text{VaR}_\alpha + E[X - \text{VaR}_\alpha | X > \text{VaR}_\alpha] \\ &= \text{VaR}_\alpha + e(\text{VaR}_\alpha - u) = \text{VaR}_\alpha + \frac{\sigma + \xi(\text{VaR}_\alpha - u)}{1 - \xi} \end{aligned}$$

## TEORIA DEI VALORI ESTREMI