

# VERJETNOST PREKOMERNEGA PRILEGNJA

Tina Janša

27. januar 2014

- Pridobitev podatkov
- Implementacija trgovalnih strategij
- Implementacija metode kombinatoričnega simetričnega navzkrižnega preverjanja

- Pridobitev podatkov
- Implementacija trgovalnih strategij
- Implementacija metode kombinatoričnega simetričnega navzkrižnega preverjanja

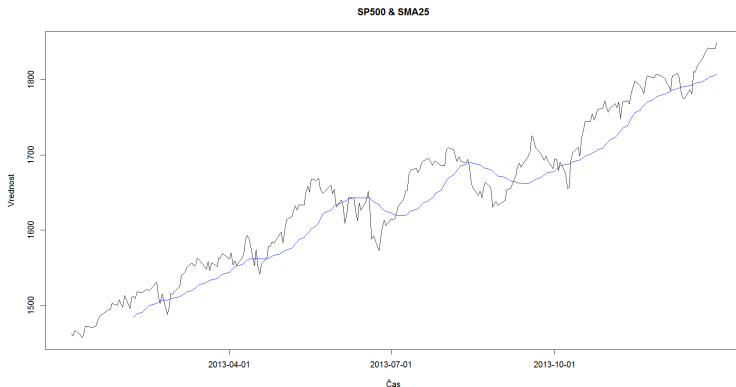
- Random
- Buy & Hold
- SMA
- RSI
- Bollinger

- Random
- Buy & Hold
- SMA
- RSI
- Bollinger

- Random
- Buy & Hold
- SMA
- RSI
- Bollinger

# SMA - Preprosto drseče povprečje

- Nakupni signal: zadnja cena  $>$  SMA
- Prodajni signal: zadnja cena  $<$  SMA

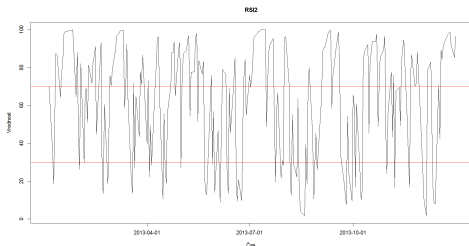


# RSI - Indeks relativne moči

- Nakupni signal:  $RSI < 30$
- Prodajni signal:  $RSI > 70$

$$RSI = 100 - \frac{100}{1 - RS}$$

$$RS = \frac{\text{povprečna vrednost pozitivnih period}}{\text{povprečna vrednost negativnih period}}$$

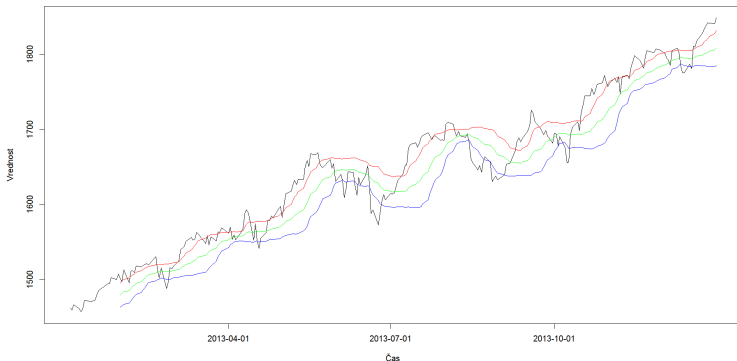




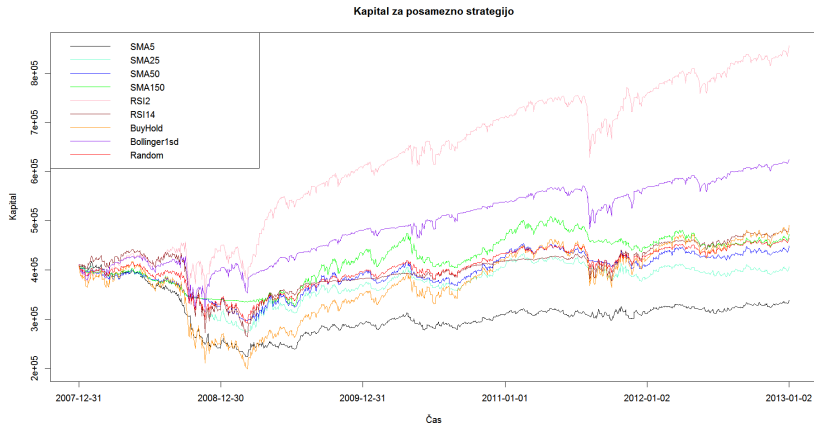
# Bollingerjevi pasovi

- Nakupni signal: zadnja cena < spodnji pas
- Prodajni signal: zadnja cena > zgornji pas
- zgornji pas =  $SMA + faktor \times standardni\ odklon$
- srednji pas = SMA
- spodnji pas =  $SMA - faktor \times standardni\ odklon$

SP500 & Bollingerjevi pasovi



# Trgovalne strategije



# Verjetnot prekomernega prileganja

- Naj bo  $R : \mathbb{R}^T \rightarrow \mathbb{R}$  funkcija, ki za dano časovno vrsto uspešnosti  $n$ -te strategije (velikosti  $T$ ) vrne realno mero uspešnosti  $R_n$ .
- Za dano množico  $N$  strategij, naj bo  $\Omega$  vzorčni prostor vseh možnih rangiranj strategij glede na  $R_n$  (v naraščajočem redu).
- Naj bo  $\mathcal{F}$  množica dogodkov, ki vsebuje sigma algebro  $\Omega$ . Natančneje, eden izmed elementov v  $\mathcal{F}$  je podmnožica vseh izidov v  $\Omega$  kjer je  $n$ -ta strategija rangirana v prvi polovici (pod mediano). Označimo ta dogodek z  $\underline{\mathcal{F}}_n \in \mathcal{F}$ ,  $\forall n = 1, \dots, N$ .
- Naj bo  $Prob$  verjetnostna mera na  $\mathcal{F}$ . Na primer, verjetnost, da je uspešnost  $n$ -te strategije pod mediano vseh uspešnosti, ustreza relativni frekvenci pri kateri se zgodi katerikoli izid v  $\underline{\mathcal{F}}_n$ .

## Prekomerno prileganje zgodovinskega preverjanja

Naj bo  $n^*$  najboljša strategija na učni množici, to je  $R_n^* \geq R_n$ , za vsak  $n = 1, \dots, N$ . Označimo z  $\overline{R}_n^*$  uspešnost strategije  $n^*$  na testni množici. Naj bo  $Me[\overline{R}]$  mediana uspešnosti vseh strategij na testni množici. Potem rečemo, da se proces izbire strategije prekomerno prilega, če za strategijo  $n^*$  z najvišjim rangom na učni množici velja:

$$E[\overline{R}_n^*] < Me[\overline{R}]$$

## Verjetnost prekomernega prileganja

Naj bo  $n^*$  najboljša strategija na učni množici. Ker  $n^*$  ni nujno najboljša strategija na testni množici, obstaja pozitivna verjetnost, da je  $\overline{R}_n^* < Me[\overline{R}]$ . Verjetnost, da se izbrana strategija  $n^*$  prekomerno prilega, definiramo kot:

$$PBO \equiv Prob[\overline{R}_n^* < Me[\overline{R}]]$$

# CSCV - Metoda kombinatoričnega simetričnega prečnega preverjanja I

1. Konstruiramo matriko  $M$  iz časovne vrste uspešnosti  $N$  strategij.  $M$  je matrika velikosti  $T \times N$ .
2.  $M$  razdelimo po vrsticah na  $S$  disjunktnih podmatrik enakih velikosti. Vsaka od teh podmatrik  $M_s$ ,  $s = 1, \dots, S$ , je velikosti  $\frac{T}{S} \times N$ .
3. Tvorimo kombinacije  $\frac{S}{2}$  podmatrik  $M_s$ , kar nam da  $\binom{S}{\frac{S}{2}}$  kombinacij  $C_s$ .
4. Za vsako kombinacijo  $c \in C_s$  naredimo naslednje:
  - Konstruiramo učno množico  $J$  tako da združimo  $\frac{S}{2}$  podmatrik  $M_s$ , ki so v kombinaciji  $c$ .  $J$  je matrika velikosti  $\frac{T}{2} \times N$ .
  - Konstruiramo testno množico  $\bar{J}$  iz podmatrik, ki niso v  $J$ .
  - Konstruiramo vektor  $R$ , kjer  $n$ -ti element pove uspešnost  $n$ -tega stolpca matrike  $J$ .

# CSCV - Metoda kombinatoričnega simetričnega prečnega preverjanja II

- Določimo element  $n^*$  tako, da je  $R_n \leq R_n^*, \forall n = 1, \dots, N$ .
  - Konstruiramo vektor  $\bar{R}$ , kjer  $n$ -ti element pove uspešnost  $n$ -tega stolpca matrike  $\bar{J}$ .
  - Določimo relativni rank  $\bar{R}_n^*$  v  $\bar{R}$ . Označimo ga z  $\bar{\omega}_c$ , kjer je  $\bar{\omega}_c \in (0, 1)$ .
  - Definiramo logit  $\lambda_c = \log \frac{\bar{\omega}_c}{1-\bar{\omega}_c}$ .  $\lambda_c = 0$ , ko je  $\bar{R}_n^* = Me[\bar{R}]$ .
- 5 Izračunamo porazdelitev rankov testne množice z zbiranjem vseh  $\lambda_c$ , za  $c \in C_s$ .  $f(\lambda)$  je relativna frekvenca pri kateri se  $\lambda$  zgodi na vseh  $C_s$ , z  $\int_{-\infty}^{\infty} f(\lambda) d\lambda = 1$ .
- 6 Verjetnost prekomernega prileganja lahko ocenimo z  $\Phi = \int_{-\infty}^0 f(\lambda) d\lambda$ .
- $\Phi \approx 0$ : ni velikega prekomernega prileganja, ker je izbor optimalne strategije na učni množici pripomogel k večji uspešnosti na testni množici.

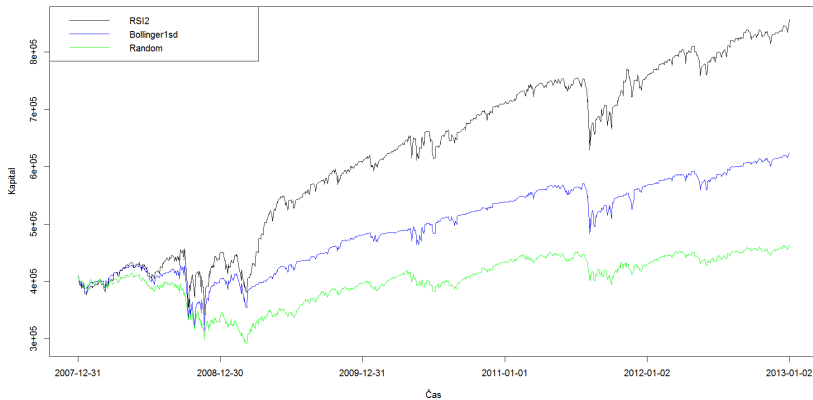
# CSCV - Metoda kombinatoričnega simetričnega prečnega preverjanja III

- $\Phi = \frac{1}{2}$ : zgodovinsko preverjanje se prekomerno prilega do te mere, da postopek izbora optimalne strategije na učni množici ne doda vrednosti.
- $\Phi \gg \frac{1}{2}$ : prekomerno prileganje je tako veliko, da izbor optimalne strategije na učni množici povzroči slabšo pričakovano uspešnost na testni množici, kot bi bila, če bi naključno izbrali eno izmed strategij.

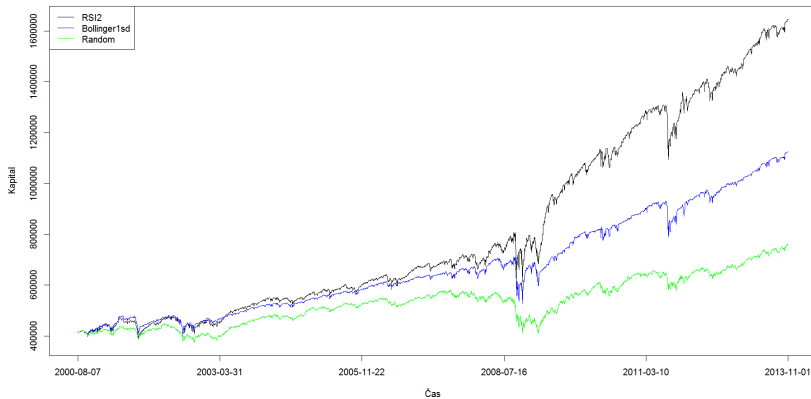
|                 |                 | STRATEGIJE                  |  |                             |   |                   |
|-----------------|-----------------|-----------------------------|--|-----------------------------|---|-------------------|
|                 |                 | SMA150<br>RSI14<br>Buy&Hold | SMA150<br>RSI14<br>Buy&Hold<br>SMA50<br>Random | RSI2<br>Bollinger<br>Random | RSI2<br>Bollinger<br>Random<br>SMA25<br>RSI14 | vse<br>strategije |
| <b>CSCV</b>     | 5 let<br>13 let | 0.7953<br>0.7416            | 0.8038<br>0.7025                               | 0.0004<br>0.0017            | 0.0026<br>0.0039                              | 0.1060<br>0.0333  |
| <b>Hold-out</b> | 5 let<br>13 let | 3<br>3                      | 5<br>5   | 3<br>3                      | 5<br>5  | vse<br>vse        |



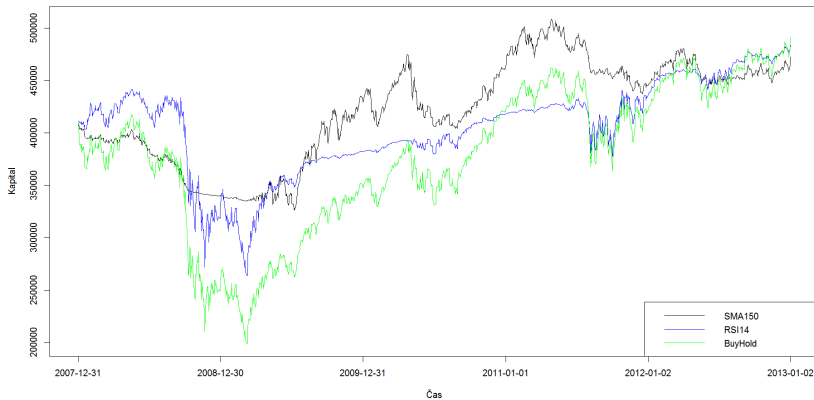
### Kapital za posamezno strategijo



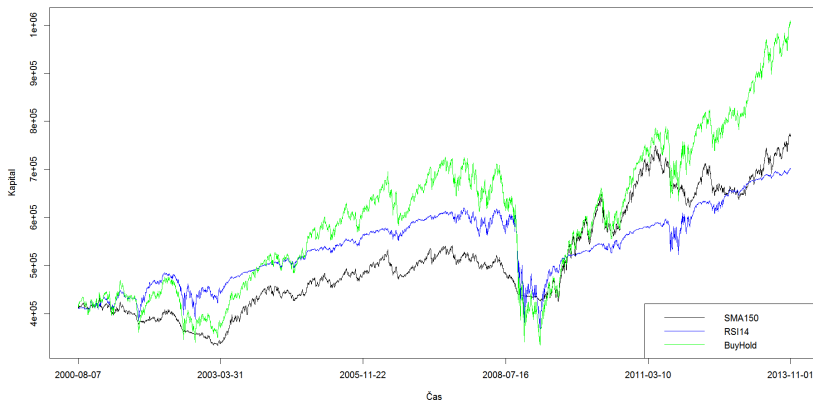
### Kapital za posamezno strategijo



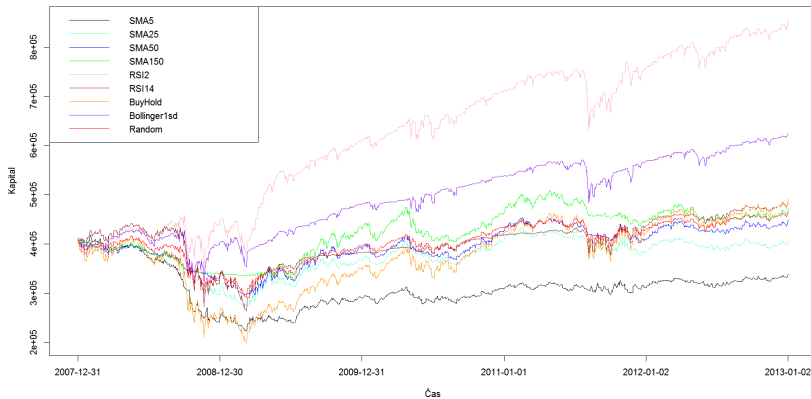
Kapital za posamezno strategijo



### Kapital za posamezno strategijo



### Kapital za posamezno strategijo



### Kapital za posamezno strategijo

