Aplikace pravděpodobnostních modelů v kurzovém sázení

Ing. Tomáš Kouřim

FJFI ČVUT

28.4.2016

Obsah

Obsah

Tenis a náhodné procesy

1. Série sportovních utkání (daného hráče, v dané soutěži) je náhodný proces
2. Průběh jednoho sportovního utkání je rovněž realizace náhodného procesu
3. Tenisový zápas v sobě obsahuje hned několik náhodných procesů
   - zápas
   - set
   - gam
   - míče

Tenis a kurzové sázení

1. Sázkařsky významný sport
   - Celosvětový impakt
   - Velké množství soutěží
   - Prakticky celoročně
   - Velký důraz na národní příslušnost
   - Individuální i týmový
2. Kurzy sázkových kanceláří jsou dobrým benchmarkem
3. V roce 2014 prosázeno v ČR 35,6 miliardy korun (na regulovaném trhu)

Obsah

Kurzové sázení

Def.: Kurz

Mějme diskrétní náhodnou veličinu X s n stavy a rozdělením p. Libovolnou n-tici $k = (k_1, \dots, k_n)$, $k_i \in \mathbb{R}$ k ní přiřazenou nazveme kurzem (kurzovou n-ticí). Sázka 1 jednotky na příležitost j znamená, že pokud je výsledkem realizace náhodné veličiny X jev j, vyplatí bookmaker sázejícímu k_j jednotek, jinak si 1 jednotku ponechá

- $k_i > 0$
- $k_i > 1$
- $k_i \in \mathbb{Q}$
- $k = f(p)$
- $k_i \approx \frac{1}{p_i}$

Kurzové sázení

Def.: Kurz

Mějme diskrétní náhodnou veličinu X s n stavy a rozdělením p. Libovolnou n-tici $k = (k_1, \ldots, k_n)$, $k_i \in \mathbb{R}$ k ní přiřazenou nazveme kurzem (kurzovou n-ticí). Sázka 1 jednotky na příležitost j znamená, že pokud je výsledkem realizace náhodné veličiny X jev j, vyplatí bookmaker sázejícímu k_j jednotek, jinak si 1 jednotku ponechá

- $k_i > 0$
- $k_i > 1$
- $k_i \in \mathbb{Q}$
- $k = f(p)$
- $k_i \approx \frac{1}{p_i}$

Kurzové sázení

Def.: Kurz

Mějme diskrétní náhodnou veličinu X s n stavy a rozdělením p. Libovolnou n-tici $k = (k_1, \ldots, k_n)$, $k_i \in \mathbb{R}$ k ní přiřazenou nazveme kurzem (kurzovou n-ticí). Sázka 1 jednotky na příležitost j znamená, že pokud je výsledkem realizace náhodné veličiny X jev j, vyplatí bookmaker sázejícímu k_j jednotek, jinak si 1 jednotku ponechá

- $k_i > 0$
- $k_i > 1$
- $k_i \in \mathbb{Q}$
- $k = f(p)$
- $k_i \approx \frac{1}{p_i}$

Kurzové sázení

Def.: Kurz

Mějme diskrétní náhodnou veličinu X s n stavy a rozdělením p. Libovolnou n-tici $k = (k_1, \ldots, k_n)$, $k_i \in \mathbb{R}$ k ní přiřazenou nazveme kurzem (kurzovou n-ticí). Sázka 1 jednotky na příležitost j znamená, že pokud je výsledkem realizace náhodné veličiny X jev j, vyplatí bookmaker sázejícímu k_j jednotek, jinak si 1 jednotku ponechá

- $k_i > 0$
- $k_i > 1$
- $k_i \in \mathbb{Q}$
- $k = f(p)$
- $k_i \approx \frac{1}{p_i}$

Kurzové sázení

Def.: Kurz

Mějme diskrétní náhodnou veličinu X s n stavy a rozdělením p. Libovolnou n-tici $k = (k_1, \ldots, k_n)$, $k_i \in \mathbb{R}$ k ní přiřazenou nazveme kurzem (kurzovou n-ticí). Sázka 1 jednotky na příležitost j znamená, že pokud je výsledkem realizace náhodné veličiny X jev j, vyplatí bookmaker sázejícímu k_j jednotek, jinak si 1 jednotku ponechá

- $k_i > 0$
- $k_i > 1$
- $k_i \in \mathbb{Q}$
- $k = f(p)$
- $k_i \approx \frac{1}{p_i}$

Kurzové sázení

Def.: Kurz

Mějme diskrétní náhodnou veličinu X s n stavy a rozdělením p. Libovolnou n-tici $k = (k_1, \ldots, k_n)$, $k_i \in \mathbb{R}$ k ní přiřazenou nazveme kurzem (kurzovou n-ticí). Sázka 1 jednotky na příležitost j znamená, že pokud je výsledkem realizace náhodné veličiny X jev j, vyplatí bookmaker sázejícímu k_j jednotek, jinak si 1 jednotku ponechá

- $k_i > 0$
- $k_i > 1$
- $k_i \in \mathbb{Q}$
- $k = f(p)$
- $k_i \approx \frac{1}{p_i}$

Kurzy a rozdělení pravděpodobnosti

- $\kappa = \sum_{i=1}^{n} \frac{1}{k_i}$
  - $\kappa = 1 \implies k$ jsou "férkurzy"
  - $\kappa > 1 \implies k$ jsou "podkurzy"
  - $\kappa < 1 \implies k$ jsou "nadkurzy"
- Pokud jsou k férkuzy, pak se je vektor $\tilde{p} = (\tilde{p}_1, \dots, , \tilde{p}_n), \tilde{p}_i = \frac{1}{k_i}$ odhadem rozdělení pravděpodobnosti p náhodné veličiny X.
- Pokud platí, že $k_i = \frac{1}{p_i} \forall i \in \hat{n}$, pak jsou k férkurzy. Opačná implikace ovšem neplatí.

Kurzy a rozdělení pravděpodobnosti

- $\kappa = \sum_{i=1}^{n} \frac{1}{k_i}$
  - $\kappa = 1 \implies k$ jsou "férkurzy"
  - $\kappa > 1 \implies k$ jsou "podkurzy"
  - $\kappa < 1 \implies k$ jsou "nadkurzy"
- Pokud jsou k férkuzy, pak se je vektor $\tilde{p} = (\tilde{p}_1, \dots, , \tilde{p}_n)$, $\tilde{p}_i = \frac{1}{k_i}$ odhadem rozdělení pravděpodobnosti p náhodné veličiny X.
- Pokud platí, že $k_i = \frac{1}{p_i}$ $\forall i \in \hat{n}$, pak jsou k férkurzy. Opačná implikace ovšem neplatí.

Kurzy a rozdělení pravděpodobnosti

- $\kappa = \sum_{i=1}^{n} \frac{1}{k_i}$
  - $\kappa = 1 \implies k$ jsou "férkurzy"
  - $\kappa > 1 \implies k$ jsou "podkurzy"
  - $\kappa < 1 \implies k$ jsou "nadkurzy"
- Pokud jsou k férkuzy, pak se je vektor $\tilde{p} = (\tilde{p}_1, \dots, , \tilde{p}_n)$, $\tilde{p}_i = \frac{1}{k_i}$ odhadem rozdělení pravděpodobnosti p náhodné veličiny X.
- Pokud platí, že $k_i = \frac{1}{p_i}\ \forall i \in \hat{n}$, pak jsou k férkurzy. Opačná implikace ovšem neplatí.

Kurzy a rozdělení pravděpodobnosti

- $\kappa = \sum_{i=1}^{n} \frac{1}{k_i}$
  - $\kappa = 1 \implies k$ jsou "férkurzy"
  - $\kappa > 1 \implies k$ jsou "podkurzy"
  - $\kappa < 1 \implies k$ jsou "nadkurzy"
- Pokud jsou k férkuzy, pak se je vektor $\tilde{p} = (\tilde{p}_1, \dots, , \tilde{p}_n)$, $\tilde{p}_i = \frac{1}{k_i}$ odhadem rozdělení pravděpodobnosti p náhodné veličiny X.
- Pokud platí, že $k_i = \frac{1}{p_i}$ $\forall i \in \hat{n}$, pak jsou k férkurzy. Opačná implikace ovšem neplatí.

Kurzy a rozdělení pravděpodobnosti

- $\kappa = \sum_{i=1}^{n} \frac{1}{k_i}$
  - $\kappa = 1 \implies k$ jsou "férkurzy"
  - $\kappa > 1 \implies k$ jsou "podkurzy"
  - $\kappa < 1 \implies k$ jsou "nadkurzy"
- Pokud jsou k férkuzy, pak se je vektor $\tilde{p} = (\tilde{p}_1, \dots, , \tilde{p}_n)$, $\tilde{p}_i = \frac{1}{k_i}$ odhadem rozdělení pravděpodobnosti p náhodné veličiny X.
- Pokud platí, že $k_i = \frac{1}{p_i}$ $\forall i \in \hat{n}$, pak jsou k férkurzy. Opačná implikace ovšem neplatí.

Kurzy a rozdělení pravděpodobnosti

- $\kappa = \sum_{i=1}^{n} \frac{1}{k_i}$
  - $\kappa = 1 \implies k$ jsou "férkurzy"
  - $\kappa > 1 \implies k$ jsou "podkurzy"
  - $\kappa < 1 \implies k$ jsou "nadkurzy"
- Pokud jsou k férkuzy, pak se je vektor $\tilde{p} = (\tilde{p}_1, \dots, , \tilde{p}_n)$, $\tilde{p}_i = \frac{1}{k_i}$ odhadem rozdělení pravděpodobnosti p náhodné veličiny X.
- Pokud platí, že $k_i = \frac{1}{p_i}\ \forall i \in \hat{n}$, pak jsou k férkurzy. Opačná implikace ovšem neplatí.

Kurzy a rozdělení pravděpodobnosti II

- Libovolné kurzy k lze převést na férkuzy normalizací
- $f(k_i) = \tilde{p}_i = \frac{1}{k_i \cdot \kappa}$ je odhadem pravděpodobnosti p_i

Příklad

$k = (1.01, 9)$

$p_k = (0.99, 0.111)$

$\tilde{p}_i = p_{k_i} - p_{k_i} \cdot (1 - payout)$

$k_{real} = (1.01, 40)$

$\kappa = 1.101$

payout $\ldots \frac{1}{1.101} = 0.9$

$\tilde{p} = (0.9, 0.1)$

$p_{real} = (0.975, 0.025)$

Kurzy a rozdělení pravděpodobnosti II

- Libovolné kurzy k lze převést na férkuzy normalizací
- $f(k_i) = \tilde{p}_i = \frac{1}{k_i \cdot \kappa}$ je odhadem pravděpodobnosti p_i

Příklad

$k = (1.01, 9)$

$p_k = (0.99, 0.111)$

$\tilde{p}_i = p_{k_i} - p_{k_i} \cdot (1 - payout)$

$k_{real} = (1.01, 40)$

$\kappa = 1.101$

payout $\ldots \frac{1}{1.101} = 0.9$

$\tilde{p} = (0.9, 0.1)$

$p_{real} = (0.975, 0.025)$

Kurzy a rozdělení pravděpodobnosti II

- Libovolné kurzy k lze převést na férkuzy normalizací
- $f(k_i) = \tilde{p}_i = \frac{1}{k_i \cdot \kappa}$ je odhadem pravděpodobnosti p_i

Příklad

$$k = (1.01, 9)$$
$$p_k = (0.99, 0.111)$$
$$\tilde{p}_i = p_{k_i} - p_{k_i} \cdot (1 - payout)$$
$$k_{real} = (1.01, 40)$$

$$\kappa = 1.101$$
$$\text{payout} \ldots \frac{1}{1.101} = 0.9$$
$$\tilde{p} = (0.9, 0.1)$$
$$p_{real} = (0.975, 0.025)$$

Normalizace kurzů

- Opačný přístup
  - marže u outsidera
  - kurz na favorita blízký férkurzu
- $g(k_i) = \tilde{p}_i = \frac{k_i(n-1)}{(n-1)+k_i(1-\frac{1}{f(k_i)})(\kappa-1)}$.

Příklad

$k = (1.01, 9)$ $\qquad$ $\kappa = 1.101$

$\tilde{p}_i = p_{k_i} - \mathring{p}_{k_i} \cdot (1 - payout)$ $\qquad$ $\tilde{p} = (0.98, 0.02)$

$k_{real} = (1.01, 40)$ $\qquad$ $p_{real} = (0.975, 0.025)$

- Lineární interpolace mezi uvedenými dvěma přístupy.

Normalizace kurzů

- Opačný přístup
  - marže u outsidera
  - kurz na favorita blízký férkurzu
- $g(k_i) = \tilde{p}_i = \frac{k_i(n-1)}{(n-1)+k_i(1-\frac{1}{f(k_i)})(\kappa-1)}.$

Příklad

$k = (1.01, 9)$ $\kappa = 1.101$

$\tilde{p}_i = p_{k_i} - \mathring{p}_{k_i} \cdot (1 - payout)$ $\tilde{p} = (0.98, 0.02)$

$k_{real} = (1.01, 40)$ $p_{real} = (0.975, 0.025)$

- Lineární interpolace mezi uvedenými dvěma přístupy.

Normalizace kurzů

- Opačný přístup
  - marže u outsidera
  - kurz na favorita blízký férkurzu
- $g(k_i) = \tilde{p}_i = \frac{k_i(n-1)}{(n-1)+k_i(1-\frac{1}{f(k_i)})(\kappa-1)}$.

Příklad

$$k = (1.01, 9) \qquad \kappa = 1.101$$

$$\tilde{p}_i = p_{k_i} - \mathring{p}_{k_i} \cdot (1 - payout) \qquad \tilde{p} = (0.98, 0.02)$$

$$k_{real} = (1.01, 40) \qquad p_{real} = (0.975, 0.025)$$

- Lineární interpolace mezi uvedenými dvěma přístupy.

Obsah

Posloupnost alternativních rozdělení

- Každé tenisové utkání je unikátní
  - Různí hráči
  - Různé podmínky
  - Různé období
- Máme vždy jen jedno pozorování dané náhodné veličiny X_j
- Každá má alternativní rozdělení s parametrem p_j
- Lyapunov CLT $\implies \sum_{j=1}^{n} X_i \sim N(\sum p_j, \sum p_j \cdot (1 - p_j))$

Posloupnost alternativních rozdělení

- Každé tenisové utkání je unikátní
  - Různí hráči
  - Různé podmínky
  - Různé období
- Máme vždy jen jedno pozorování dané náhodné veličiny X_j
- Každá má alternativní rozdělení s parametrem p_j
- Lyapunov CLT $\implies \sum_{j=1}^{n} X_i \sim N(\sum p_j, \sum p_j \cdot (1 - p_j))$

Posloupnost alternativních rozdělení

- Každé tenisové utkání je unikátní
  - Různí hráči
  - Různé podmínky
  - Různé období
- Máme vždy jen jedno pozorování dané náhodné veličiny X_j
- Každá má alternativní rozdělení s parametrem p_j
- Lyapunov CLT $\implies \sum_{j=1}^{n} X_i \sim N(\sum p_j, \sum p_j \cdot (1 - p_j))$

Posloupnost alternativních rozdělení

- Každé tenisové utkání je unikátní
  - Různí hráči
  - Různé podmínky
  - Různé období
- Máme vždy jen jedno pozorování dané náhodné veličiny X_j
- Každá má alternativní rozdělení s parametrem p_j
- Lyapunov CLT $\implies \sum_{j=1}^{n} X_i \sim N(\sum p_j, \sum p_j \cdot (1 - p_j))$

Sázecí strategie

- Místo počtu výher hráče můžeme sledovat celkovou bilanci "sázkové strategie"
  1. Sázka 1 jednotka
  2. Sázka $\frac{1}{k}$
  3. p, $k-1$, ...
- Každá strategie konverguje k určitému normálnímu rozdělení
- Můžeme provádět standardní testování hypotéz

Sázecí strategie

- Místo počtu výher hráče můžeme sledovat celkovou bilanci "sázkové strategie"
  1. Sázka 1 jednotka
  2. Sázka $\frac{1}{k}$
  3. p, $k-1$, ...
- Každá strategie konverguje k určitému normálnímu rozdělení
- Můžeme provádět standardní testování hypotéz

Sázecí strategie

- Místo počtu výher hráče můžeme sledovat celkovou bilanci "sázkové strategie"
  1. Sázka 1 jednotka
  2. Sázka $\frac{1}{k}$
  3. p, $k-1$, ...
- Každá strategie konverguje k určitému normálnímu rozdělení
- Můžeme provádět standardní testování hypotéz

Sázecí strategie

- Místo počtu výher hráče můžeme sledovat celkovou bilanci "sázkové strategie"
  1. Sázka 1 jednotka
  2. Sázka $\frac{1}{k}$
  3. p, $k-1$, ...
- Každá strategie konverguje k určitému normálnímu rozdělení
- Můžeme provádět standardní testování hypotéz

Sázecí strategie

- Místo počtu výher hráče můžeme sledovat celkovou bilanci "sázkové strategie"
  1. Sázka 1 jednotka
  2. Sázka $\frac{1}{k}$
  3. p, $k-1$, ...
- Každá strategie konverguje k určitému normálnímu rozdělení
- Můžeme provádět standardní testování hypotéz

Sázecí strategie

- Místo počtu výher hráče můžeme sledovat celkovou bilanci "sázkové strategie"
  1. Sázka 1 jednotka
  2. Sázka $\frac{1}{k}$
  3. p, $k-1$, ...
- Každá strategie konverguje k určitému normálnímu rozdělení
- Můžeme provádět standardní testování hypotéz

Porovnání strategií

- Předpokládáme, že se vždy jedná o férkurzy

Strategie "1"

$$E(X) = 0$$
$$Var(X) = \sum_{j=1}^{n} \frac{1-p_j}{p_j}$$
$$(X_{h1} + X_{h2})\ ?\ 0$$
$$Var(X_{h1}) \neq Var(X_{h2})$$

Strategie "1/k"

$$E(X) = 0$$
$$Var(X) = \sum_{j=1}^{n} p_j(1-p_j)$$
$$(X_{h1} + X_{h2}) = 0$$
$$Var(X_{h1}) = Var(X_{h2})$$

Porovnání strategií

- Předpokládáme, že se vždy jedná o férkurzy

Strategie "1"

$$E(X) = 0$$

$$Var(X) = \sum_{j=1}^{n} \frac{1-p_j}{p_j}$$

$$(X_{h1} + X_{h2}) \; ? \; 0$$

$$Var(X_{h1}) \neq Var(X_{h2})$$

Strategie "1/k"

$$E(X) = 0$$

$$Var(X) = \sum_{j=1}^{n} p_j(1-p_j)$$

$$(X_{h1} + X_{h2}) = 0$$

$$Var(X_{h1}) = Var(X_{h2})$$

Příklad

k_1	k_2	p_1	p_2	h_1	h_2	$zisk_1$	$zisk_2$
1.25	5	0.8	0.2	1	0	0.25	-1
1.5	3	0.66	0.33	1	0	0.5	-1
1.8	2.25	0.55	0.44	1	0	0.8	-1
1.01	101	0.99	0.01	0	1	-1	100

Table: Příklad aplikace sázkové strategie "1".

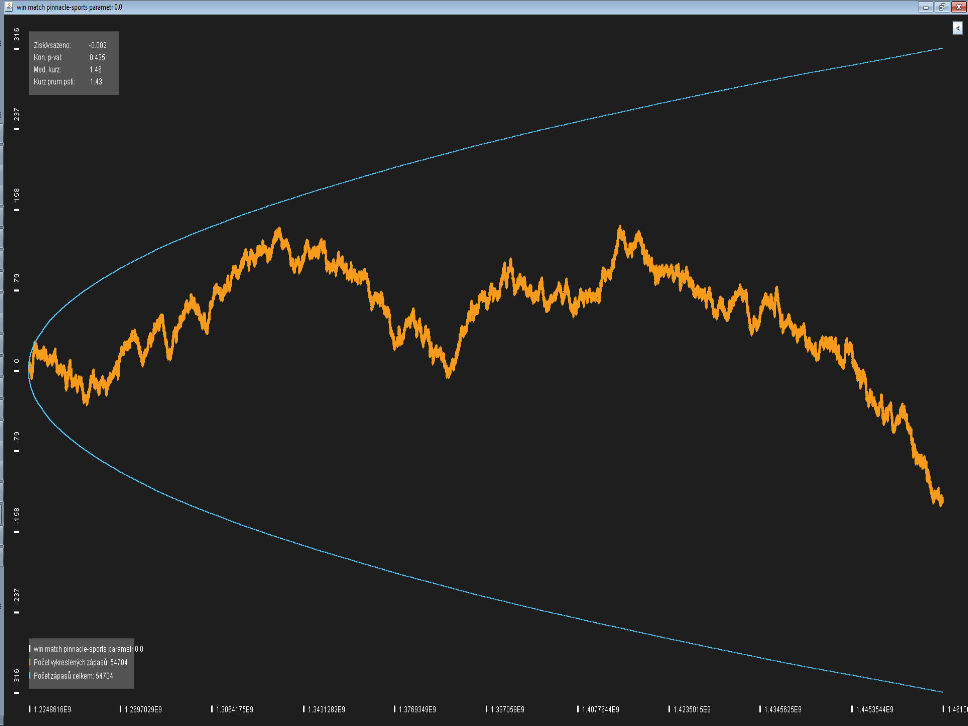
Příklad

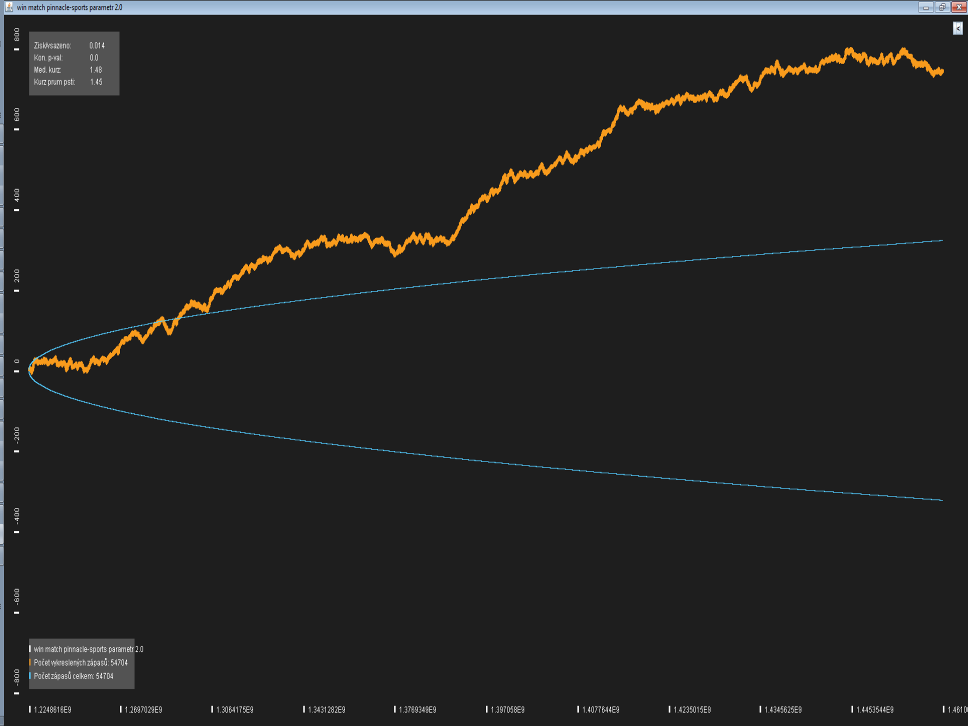
k_1	k_2	p_1	p_2	h_1	h_2	$zisk_1$	$zisk_2$
1.25	5	0.8	0.2	1	0	0.2	-0.2
1.5	3	0.66	0.33	1	0	0.33	-0.33
1.8	2.25	0.55	0.44	1	0	0.44	-0.44
1.01	101	0.99	0.01	0	1	-0.99	0.99

Table: Příklad aplikace sázkové strategie "1/k".

Další kritéria hodnocení

1. Log-likelihood
2. Parametry jako extrémní či průměrné ochylky od očekávaných hodnot
3. Vše lze řešit i pro nejrůznější podskupiny pozorování, ideálně pro všechny (výpočetně nemožné)


win match pinnacle-sports parametr 0.0
Zisk/vsazeno: -0.002
Kon. p-val: 0.435
Med. kurz: 1.46
Kurz prum psti: 1.43
316
237
158
79
0
-79
-158
-237
-316
win match pinnacle-sports parametr 0.0
Počet vykreslených zápasů: 54704
Počet zápasů celkem: 54704
1.2248616E9
1.2697029E9
1.3064175E9
1.3431282E9
1.3769349E9
1.397058E9
1.4077644E9
1.4235015E9
1.4345625E9
1.4453544E9


win match pinnacle-sports parametr 2.0
Zisk/vsazeno: 0.014
Kon. p-val: 0.0
Med. kurz: 1.48
Kurz prum psti: 1.45
800
600
400
200
0
-200
-400
-600
-800
win match pinnacle-sports parametr 2.0
Počet vykreslených zápasů: 54704
Počet zápasů celkem: 54704
1.2248616E9
1.2697029E9
1.3064175E9
1.3431282E9
1.3769349E9
1.397058E9
1.4077644E9
1.4235015E9
1.4345625E9
1.4453544E9
1.4610


win match pinnacle-sports parametr 0.0
Zisk/vsazeno: -0.002
Kon. p-val: 0.435
Med. kurz: 1.46
Kurz prum psti: 1.43
316
237
158
79
0
-79
-158
-237
-316
win match pinnacle-sports parametr 0.0
Počet vykreslených zápasů: 54704
Počet zápasů celkem: 54704
0.5
0.54
0.58
0.61
0.65
0.68
0.72
0.76
0.81
0.88
1.0


win match pinnacle-sports parametr 2.0
Zisk/vsazeno: 0.014
Kon. p-val: 0.0
Med. kurz: 1.48
Kurz prum psti: 1.45
746
559
373
186
0
-186
-373
-559
-746
win match pinnacle-sports parametr 2.0
Počet vykreslených zápasů: 54704
Počet zápasů celkem: 54704
0.5
0.54
0.58
0.61
0.65
0.68
0.72
0.76
0.81
0.88
1.0

Obsah

Predikce na základě historických výsledků

- Přímočarý přístup
- Cílem je co nejpřesněji odhadnout vítěze
- Zkreslená a neúplná data
  - Výsledek ovlivňují objektivní ale i subjektivní data
  - Subjektivní data jsou těžko kvantifikovaltelná
  - A ještě hůře dohledatelná
- Vyřešený problém

Predikce na základě historických výsledků

- Přímočarý přístup
- Cílem je co nejpřesněji odhadnout vítěze
- Zkreslená a neúplná data
  - Výsledek ovlivňují objektivní ale i subjektivní data
  - Subjektivní data jsou těžko kvantifikovaltelná
  - A ještě hůře dohledatelná
- Vyřešený problém

Predikce na základě historických výsledků

- Přímočarý přístup
- Cílem je co nejpřesněji odhadnout vítěze
- Zkreslená a neúplná data
  - Výsledek ovlivňují objektivní ale i subjektivní data
  - Subjektivní data jsou těžko kvantifikovaltelná
  - A ještě hůře dohledatelná
- Vyřešený problém

Predikce na základě historických výsledků

- Přímočarý přístup
- Cílem je co nejpřesněji odhadnout vítěze
- Zkreslená a neúplná data
  - Výsledek ovlivňují objektivní ale i subjektivní data
  - Subjektivní data jsou těžko kvantifikovaltelná
  - A ještě hůře dohledatelná
- Vyřešený problém

Predikce na základě historických výsledků

- Přímočarý přístup
- Cílem je co nejpřesněji odhadnout vítěze
- Zkreslená a neúplná data
  - Výsledek ovlivňují objektivní ale i subjektivní data
  - Subjektivní data jsou těžko kvantifikovaltelná
  - A ještě hůře dohledatelná
- Vyřešený problém

Predikce na základě historických kurzů

- Cílem je odhadnout dílčí výsledek utkání
- Anomálie lépe podchycené ve vstupních datech
- Vstupní data jsou relativně dobře dohledatelné a to včetně historie
- Více prostoru pro predikci - teoreticky až na úroveň jednotlivých míčků
- "Díry na trhu"

Obsah

Markovské procesy

- Pravděpodobnost výskytu v čase t závisí pouze na pozici ve stavu $t-1$
- Bezpaměťový náhodný proces - Markovská vlastnost
- Diskrétní množina stavů
- Dělí se na konečné, nekonečné, homogenní, nehomogenní,...
- Tenisové utkání je konečný absorbční Markovův proces (nebo může být)

Markovský proces

- Je popsán
  1. vektorem počátečního rozdělení $p(0) = (p_1, \ldots, p_N)$
  2. maticí přechodu P
- $p_i(n)$ představuje pravděpodobnost, že se proces bude v kroku n nacházet ve stavu j.
- p_{ij} je pravděpodobnost, že se proces dostane z aktuálního stavu i v příštím kroce do stavu j.
- $p(n) = p(0) \cdot \mathrm{P}^n$
- Existují jednoduché nástroje pro výpočet celé řady zajímavých aspektů Markovského řetězce.

Markovský proces

- Je popsán
  1. vektorem počátečního rozdělení $p(0) = (p_1, \ldots, p_N)$
  2. maticí přechodu P
- $p_i(n)$ představuje pravděpodobnost, že se proces bude v kroku n nacházet ve stavu j.
- p_{ij} je pravděpodobnost, že se proces dostane z aktuálního stavu i v příštím kroce do stavu j.
- $p(n) = p(0) \cdot \mathrm{P}^n$
- Existují jednoduché nástroje pro výpočet celé řady zajímavých aspektů Markovského řetězce.

Markovský proces

- Je popsán
  1. vektorem počátečního rozdělení $p(0) = (p_1, \ldots, p_N)$
  2. maticí přechodu P
- $p_i(n)$ představuje pravděpodobnost, že se proces bude v kroku n nacházet ve stavu j.
- p_{ij} je pravděpodobnost, že se proces dostane z aktuálního stavu i v příštím kroce do stavu j.
- $p(n) = p(0) \cdot \mathrm{P}^n$
- Existují jednoduché nástroje pro výpočet celé řady zajímavých aspektů Markovského řetězce.

Markovský proces

- Je popsán
  1. vektorem počátečního rozdělení $p(0) = (p_1, \dots, p_N)$
  2. maticí přechodu P
- $p_i(n)$ představuje pravděpodobnost, že se proces bude v kroku n nacházet ve stavu j.
- p_{ij} je pravděpodobnost, že se proces dostane z aktuálního stavu i v příštím kroce do stavu j.
- $p(n) = p(0) \cdot \mathrm{P}^n$
- Existují jednoduché nástroje pro výpočet celé řady zajímavých aspektů Markovského řetězce.

Obsah

Markovská vlastnost v tenise - sety

- Sleduji vývoj utkání hraného na dva vítězné sety
- Proces má 4 tranzientní stavy a dva absorbční
- Neměnná pravděpodobnost výhry setu
  - Pravděpodobnosti přechodu jsou stejné pro všechny tranzientní stavy, tedy jsou i.i.d.
  - Řešení rovnice $p_{zapasu} = p_{setu}^2 + 2 \cdot p_{setu}^2 \cdot (1 - p_{setu})$
- Pravděpodobnost závislá pouze na stavu.
  - Stav 1:1 má předchodové pravděpodobnosti nezávislé na cestě do tohoto stavu.
- Výsledky prvního setu odpovídají markovskému předpokladu s neměnnou pravděpodobností výhry v setu.

p_z lower	p_z upper	p_s i.i.d.	2nd set win ratio	total	p_value
0,50	0,53	0,51	0,67	226	0,000
0,53	0,57	0,54	0,63	813	0,000
0,57	0,61	0,56	0,65	857	0,000
0,61	0,66	0,59	0,71	892	0,000
0,66	0,70	0,61	0,70	1257	0,000
0,70	0,76	0,66	0,74	1589	0,000
0,76	0,81	0,70	0,78	998	0,000
0,81	0,87	0,75	0,80	1117	0,000
0,87	0,93	0,80	0,87	930	0,000
0,93	1	0,88	0,89	634	0,168
0,5	1	0,65	0,75	9313	0,000

Table: The favorite has won the first set.

p_z lower	p_z upper	p_s i.i.d.	2nd set win ratio	total	p_value
0,50	0,53	0,51	0,36	216	0,000
0,53	0,57	0,54	0,41	713	0,000
0,57	0,61	0,56	0,46	644	0,000
0,61	0,66	0,59	0,48	669	0,000
0,66	0,70	0,61	0,50	821	0,000
0,70	0,76	0,66	0,55	759	0,000
0,76	0,81	0,70	0,60	428	0,000
0,81	0,87	0,75	0,59	370	0,000
0,87	0,93	0,80	0,67	243	0,000
0,93	1	0,88	0,75	64	0,001
0,5	1	0,65	0,51	4927	0,000

Table: The favorite has lost the first set.

Markovská vlastnost v tenise - sety

- Sleduji vývoj utkání hraného na dva vítězné sety
- Proces má 4 tranzientní stavy a dva absorbční
- Neměnná pravděpodobnost výhry setu
  - Pravděpodobnosti přechodu jsou stejné pro všechny tranzientní stavy, tedy jsou i.i.d.
  - Řešení rovnice $p_{zapasu} = p_{setu}^2 + 2 \cdot p_{setu}^2 \cdot (1 - p_{setu})$
- Pravděpodobnost závislá pouze na stavu.
  - Stav 1:1 má předchodové pravděpodobnosti nezávislé na cestě do tohoto stavu.
- Výsledky prvního setu odpovídají markovskému předpokladu s neměnnou pravděpodobností výhry v setu.

p_z lower	p_z upper	outs fav	total	fav outs	total	p-value
0,50	0,53	0,65	78	0,47	75	0,009
0,53	0,57	0,58	291	0,41	300	0,000
0,57	0,61	0,59	294	0,50	300	0,012
0,61	0,66	0,56	320	0,55	260	0,375
0,66	0,70	0,63	412	0,59	380	0,130
0,70	0,76	0,66	414	0,65	416	0,404
0,76	0,81	0,73	255	0,73	222	0,421
0,81	0,87	0,76	217	0,75	228	0,444
0,87	0,93	0,79	163	0,76	118	0,285
0,93	1	0,83	48	0,81	70	0,394
0,5	1	0,65	2492	0,60	2369	0,000

Table: Comparison of the two ways of getting into the 1:1 state of a tennis match.

Markovská vlastnost v tenise - sety

- Sleduji vývoj utkání hraného na dva vítězné sety
- Proces má 4 tranzientní stavy a dva absorbční
- Neměnná pravděpodobnost výhry setu
  - Pravděpodobnosti přechodu jsou stejné pro všechny tranzientní stavy, tedy jsou i.i.d.
  - Řešení rovnice $p_{zapasu} = p_{setu}^2 + 2 \cdot p_{setu}^2 \cdot (1 - p_{setu})$
- Pravděpodobnost závislá pouze na stavu.
  - Stav 1:1 má předchodové pravděpodobnosti nezávislé na cestě do tohoto stavu.
- Výsledky prvního setu odpovídají markovskému předpokladu s neměnnou pravděpodobností výhry v setu.

p_z lower	p_z upper	p_s i.i.d.	1st set win ratio	total	p_value
0,50	0,53	0,51	0,51	442	0,491
0,53	0,57	0,54	0,53	1526	0,398
0,57	0,61	0,56	0,57	1501	0,264
0,61	0,66	0,59	0,57	1561	0,123
0,66	0,70	0,61	0,60	2078	0,087
0,70	0,76	0,66	0,68	2348	0,032
0,76	0,81	0,70	0,70	1426	0,361
0,81	0,87	0,75	0,75	1487	0,429
0,87	0,93	0,80	0,79	1173	0,178
0,93	1	0,88	0,91	698	0,006
0,5	1	0,65	0,65	14240	0,476

Table: The i.i.d. hypothesis compared to the actual results of first sets.

Obsah

Vstupní data

- Výsledky zápasů ATP z let 2009 až 2014.
  - Pouze turnaje hrané na 2 vítězné sety.
  - Turnaje série Challenger, ITF a veškeré ženské turnaje zatím vynechány.
- Historické kurzy od devíti sázkových kanceláří.
  - Tipsport, Fortuna, bet365, bwin, bet-at-home, Interwetten, Sportingbet, Unibet, William Hill.
  - navíc průměr a maximum.
- Data získána ze serveru www.livesport.cz

Použití dat

1. Skutečné kurzy sázkových kanceláří na vítěze zápasu převedu na férkurzy.
2. Jako vstup pro výpočty vyberu tu sadu kurzů, která nejlépe odpovídá skutečnosti (bet365).
3. Pravděpodobnosti výhry v prvním setu získám řešením rovnice $p_{zapasu} = p_{setu}^2 + 2 \cdot p_{setu}^2 \cdot (1 - p_{setu})$.
4. Porovnám získané pravděpodobnosti s kurzy vypsanými jednotlivými sázkovými kancelářemi.

Použití dat

1. Skutečné kurzy sázkových kanceláří na vítěze zápasu převedu na férkurzy.
2. Jako vstup pro výpočty vyberu tu sadu kurzů, která nejlépe odpovídá skutečnosti (bet365).
3. Pravděpodobnosti výhry v prvním setu získám řešením rovnice $p_{zapasu} = p_{setu}^2 + 2 \cdot p_{setu}^2 \cdot (1 - p_{setu})$.
4. Porovnám získané pravděpodobnosti s kurzy vypsanými jednotlivými sázkovými kancelářemi.

Použití dat

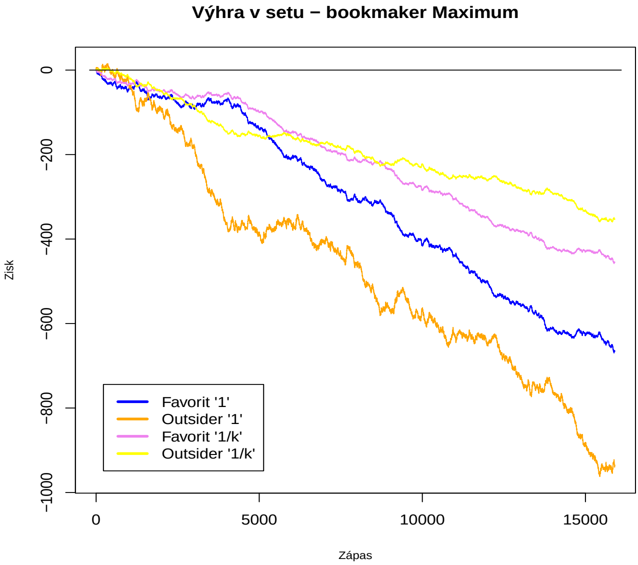
1. Skutečné kurzy sázkových kanceláří na vítěze zápasu převedu na férkurzy.
2. Jako vstup pro výpočty vyberu tu sadu kurzů, která nejlépe odpovídá skutečnosti (bet365).
3. Pravděpodobnosti výhry v prvním setu získám řešením rovnice $p_{zapasu} = p_{setu}^2 + 2 \cdot p_{setu}^2 \cdot (1 - p_{setu})$.
4. Porovnám získané pravděpodobnosti s kurzy vypsanými jednotlivými sázkovými kancelářemi.

Použití dat

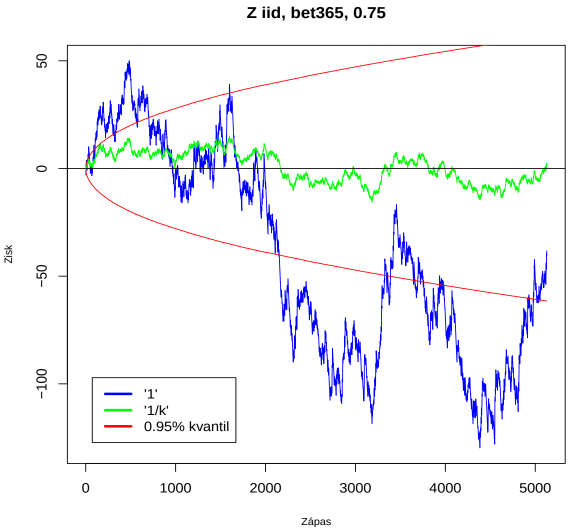
1. Skutečné kurzy sázkových kanceláří na vítěze zápasu převedu na férkurzy.
2. Jako vstup pro výpočty vyberu tu sadu kurzů, která nejlépe odpovídá skutečnosti (bet365).
3. Pravděpodobnosti výhry v prvním setu získám řešením rovnice $p_{zapasu} = p_{setu}^2 + 2 \cdot p_{setu}^2 \cdot (1 - p_{setu})$.
4. Porovnám získané pravděpodobnosti s kurzy vypsanými jednotlivými sázkovými kancelářemi.

Obsah

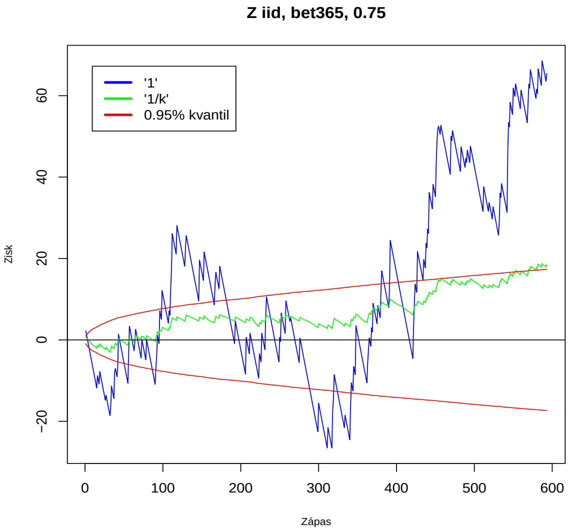
Sázení na všechna utkání



Sázení na vyšší kurz



Sázení na o 10 % vyšší kurz



Přehled výsledků sázení

Strategie	Zisk	ROI	Sázek	Min	ROI 2
Favorit "1"	-666.21	-4.19 %	15 898	-669.78	-99.4 %
Favorit "$1/k$"	-455.57	-4.18 %	15 898	-458.1	-99.4 %
Vyšší kurz "1"	-38.24	-0.75 %	5128	-129.76	-29.47 %
Vyšší kurz "$1/k$"	2.5	0.15 %	5128	-15.72	15.9 %
10% vyšší "1"	65.47	11.04 %	593	-26.54	247 %
10% vyšší "$1/k$"	18.59	21.77 %	593	-3.02	615 %

Závěr

- Úvod do kurzového sázení
- Ukázka jednoduchého modelu a jeho výsledků
- Sázení je gambling!

Závěr

- Úvod do kurzového sázení
- Ukázka jednoduchého modelu a jeho výsledků
- Sázení je gambling!

Děkuji za pozornost.

tom@skourim.com