Podruhé

o výpočtu Blakerova konfidenčního intervalu:

Balíček `BlakerCI` a jiné resty

Jan Klaschka
klaschka@cs.cas.cz
ÚI AV ČR, Praha

ROBUST 2012, Němčičky 9. – 14. září 2012

Osnova

- Pro ty, kdo na ROBUSTu 2010 chyběli nebo nedávali pozor:

Osnova

- Pro ty, kdo na ROBUSTu 2010 chyběli nebo nedávali pozor:
  - Co jest Blakerův konfidenční interval (pro parametr binomického rozdělení)

Osnova

- Pro ty, kdo na ROBUSTu 2010 chyběli nebo nedávali pozor:
  - Co jest Blakerův konfidenční interval (pro parametr binomického rozdělení)
  - Původní Blakerův algoritmus a jeho nedostatky

Osnova

- Pro ty, kdo na ROBUSTu 2010 chyběli nebo nedávali pozor:
  - Co jest Blakerův konfidenční interval (pro parametr binomického rozdělení)
  - Původní Blakerův algoritmus a jeho nedostatky
  - Můj algoritmus a jeho přednosti

Osnova

- Pro ty, kdo na ROBUSTu 2010 chyběli nebo nedávali pozor:
  - Co jest Blakerův konfidenční interval (pro parametr binomického rozdělení)
  - Původní Blakerův algoritmus a jeho nedostatky
  - Můj algoritmus a jeho přednosti
- Balíček `BlakerCI` v R

Osnova

- Pro ty, kdo na ROBUSTu 2010 chyběli nebo nedávali pozor:
  - Co jest Blakerův konfidenční interval (pro parametr binomického rozdělení)
  - Původní Blakerův algoritmus a jeho nedostatky
  - Můj algoritmus a jeho přednosti
- Balíček `BlakerCI` v R
- Konkurenční algoritmus M. P. Faye (a jeho přednosti i nedostatky)

Osnova

- Pro ty, kdo na ROBUSTu 2010 chyběli nebo nedávali pozor:
  - Co jest Blakerův konfidenční interval (pro parametr binomického rozdělení)
  - Původní Blakerův algoritmus a jeho nedostatky
  - Můj algoritmus a jeho přednosti
- Balíček `BlakerCI` v R
- Konkurenční algoritmus M. P. Faye (a jeho přednosti i nedostatky)
- O radostech, jež skýtá počítačová aritmetika s omezenou přesností

Blakerův confidenční interval

- Úloha: k . . . realizace $X \sim Binom(n, p)$, n pevné, známé $\rightarrow$ exaktní (konzervativní) dvoustranný konfidenční interval (CI) pro p na hladině spolehlivosti $1 - \alpha$

Blakerův confidenční interval

- Úloha: $k \ldots$ realizace $X \sim Binom(n, p)$, n pevné, známé $\rightarrow$ exaktní (konzervativní) dvoustranný konfidenční interval (CI) pro p na hladině spolehlivosti $1 - \alpha$

- Clopper-Pearson (1934): Hlídá se současně

  $$P_p(\text{celý CI je pod skutečným } p) \leq \alpha/2,$$

  $$P_p(\text{celý CI je nad skutečným } p) \leq \alpha/2$$

Blakerův confidenční interval

- Úloha: k ... realizace $X \sim Binom(n, p)$, n pevné, známé $\rightarrow$ exaktní (konzervativní) dvoustranný konfidenční interval (CI) pro p na hladině spolehlivosti $1 - \alpha$
- Clopper-Pearson (1934): Hlídá se současně

  P_p(celý CI je pod skutečným p) $\leq \alpha/2$,

  P_p(celý CI je nad skutečným p) $\leq \alpha/2$
- Méně konzervativní exaktní alternativy: Hlídá se jen součet

  $P_p(\ldots \text{pod} \ldots) + P_p(\ldots \text{nad} \ldots) \leq \alpha$

Blakerův konfidenční interval

- Návrhy méně konzervativních alternativ Clopper-Pearsonova intervalu:

Blakerův konfidenční interval

- Návrhy méně konzervativních alternativ Clopper-Pearsonova intervalu:
  - Sterne, Crow (1954, 1956)

Blakerův konfidenční interval

- Návrhy méně konzervativních alternativ Clopper-Pearsonova intervalu:
  - Sterne, Crow (1954, 1956)
  - Blyth, Still, Casella (1983, 1986)

Blakerův konfidenční interval

- Návrhy méně konzervativních alternativ Clopper-Pearsonova intervalu:
  - Sterne, Crow (1954, 1956)
  - Blyth, Still, Casella (1983, 1986)
  - Blaker (2000)

Blakerův konfidenční interval

- Návrhy méně konzervativních alternativ Clopper-Pearsonova intervalu:
  - Sterne, Crow (1954, 1956)
  - Blyth, Still, Casella (1983, 1986)
  - Blaker (2000)
- Přednosti Blakerova CI:

Blakerův konfidenční interval

- Návrhy méně konzervativních alternativ Clopper-Pearsonova intervalu:
  - Sterne, Crow (1954, 1956)
  - Blyth, Still, Casella (1983, 1986)
  - Blaker (2000)
- Přednosti Blakerova CI:
  - Blakerův CI $\subseteq$ Clopper-Pearsonův CI

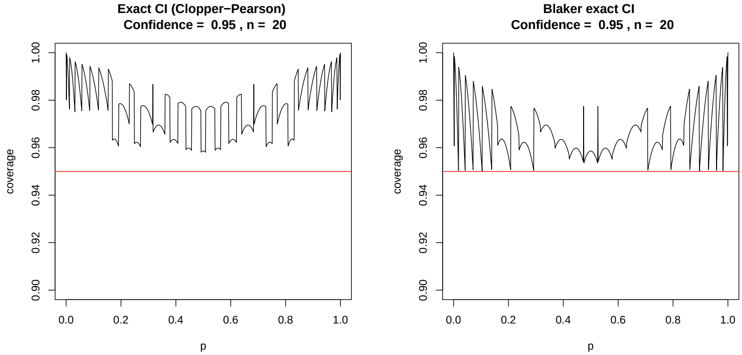
Blakerův konfidenční interval

- Návrhy méně konzervativních alternativ Clopper-Pearsonova intervalu:
  - Sterne, Crow (1954, 1956)
  - Blyth, Still, Casella (1983, 1986)
  - Blaker (2000)
- Přednosti Blakerova CI:
  - Blakerův CI $\subseteq$ Clopper-Pearsonův CI
  - Meze monotónní $\sim \alpha$

Blakerův konfidenční interval

- Návrhy méně konzervativních alternativ Clopper-Pearsonova intervalu:
  - Sterne, Crow (1954, 1956)
  - Blyth, Still, Casella (1983, 1986)
  - Blaker (2000)
- Přednosti Blakerova CI:
  - Blakerův CI $\subseteq$ Clopper-Pearsonův CI
  - Meze monotónní $\sim \alpha$
  - Snadný výpočet – v článku Blaker (2000) krátký program v R

Clopper-Pearsonův vs. Blakerův CI (pokrytí)



Blakerův konfidenční interval

- Clopper-Pearsonův CI pro $X = k$ lze psát jako

$$\{p; \beta^{CP}(p) > \alpha\},$$

kde

$$\beta^{CP}(p) = \min\{2P_p(X \geq k), 2P_p(X \leq k), 1\}$$

Blakerův konfidenční interval

- Clopper-Pearsonův CI pro $X = k$ lze psát jako

$$\{p; \beta^{CP}(p) > \alpha\},$$

kde

$$\beta^{CP}(p) = \min\{2P_p(X \geq k), 2P_p(X \leq k), 1\}$$

- Blakerův CI se definuje podobným způsobem

Blakerův konfidenční interval

- Blakerova funkce přijatelnosti (acceptability function) pro $X = k$:

$$\beta(p) = \min\{P_p(X \geq k) + P_p(X \leq k_p^*), \\ P_p(X \leq k) + P_p(X \geq k_p^{**}), 1\}$$

Blakerův konfidenční interval

- Blakerova funkce přijatelnosti (acceptability function) pro $X = k$:

$$\beta(p) = \min\{P_p(X \geq k) + P_p(X \leq k_p^*), \\ P_p(X \leq k) + P_p(X \geq k_p^{**}), 1\}$$

- k_p^* ... největší i takové, že $P_p(X \leq i) \leq P_p(X \geq k)$

Blakerův konfidenční interval

- Blakerova funkce přijatelnosti (acceptability function) pro $X = k$:

$$\beta(p) = \min\{P_p(X \geq k) + P_p(X \leq k_p^*), \\ P_p(X \leq k) + P_p(X \geq k_p^{**}), 1\}$$

- k_p^* …největší i takové, že $P_p(X \leq i) \leq P_p(X \geq k)$
- k_p^{**} …nejmenší i takové, že $P_p(X \geq i) \leq P_p(X \leq k)$

Blakerův konfidenční interval

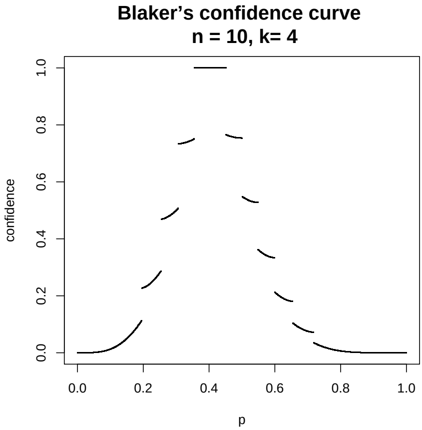
- Blakerova funkce přijatelnosti (acceptability function) pro $X = k$:

$$\beta(p) = \min\{P_p(X \geq k) + P_p(X \leq k_p^*), \\ P_p(X \leq k) + P_p(X \geq k_p^{**}), 1\}$$

- k_p^* …největší i takové, že $P_p(X \leq i) \leq P_p(X \geq k)$
- k_p^{**} …nejmenší i takové, že $P_p(X \geq i) \leq P_p(X \leq k)$
- Blakerův CI je zhruba množina $\{p; \beta(p) > \alpha\}$

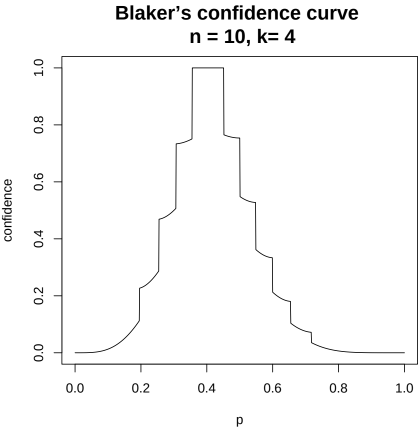
Blakerův konfidenční interval

- Příklad funkce přijatelnosti:



Blakerův konfidenční interval

- Příklad funkce přijatelnosti:

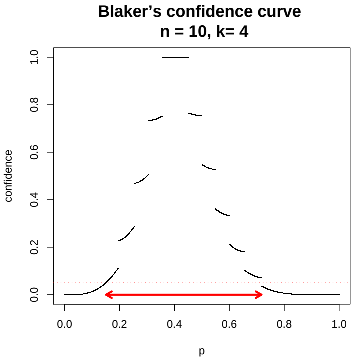


Blakerův konfidenční interval

- Zhruba: Blakerův CI na hladině spolehlivosti $(1-\alpha)$ je tam, kde $\beta(p) > \alpha$

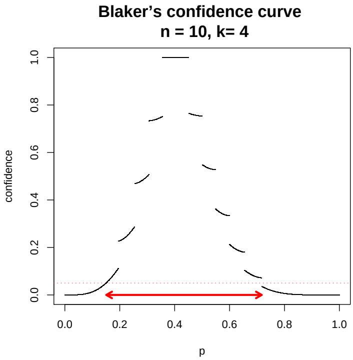
Blakerův konfidenční interval

- Zhruba: Blakerův CI na hladině spolehlivosti $(1 - \alpha)$ je tam, kde $\beta(p) > \alpha$



Blakerův konfidenční interval

- Zhruba: Blakerův CI na hladině spolehlivosti $(1-\alpha)$ je tam, kde $\beta(p) > \alpha$



- Výpočet konfidenčních mezí: Numerické hledání bodů, kde $\beta(\cdot)$ „kříží úroveň α“, tj. $\beta(\cdot)-\alpha$ mění znaménko

Blakerův konfidenční interval

- Přesněji:
  $C = \{p; \beta(p) > \alpha\}$ je vždy konfidenční množina na hladině spolehlivosti $1 - \alpha$, ale nemusí být interval (a často není)

Blakerův konfidenční interval

- Přesněji:
  $C = \{p; \beta(p) > \alpha\}$ je vždy konfidenční množina na hladině spolehlivosti $1 - \alpha$, ale nemusí být interval (a často není)
- Blakerův interval je třeba definovat jako $conv(C)$ (konvexní obal)

Blakerův konfidenční interval

- Přesněji:
  $C = \{p; \beta(p) > \alpha\}$ je vždy konfidenční množina na hladině spolehlivosti $1 - \alpha$, ale nemusí být interval (a často není)
- Blakerův interval je třeba definovat jako $conv(C)$ (konvexní obal)
- Výpočet konfidenčních mezí:
  Numerické hledání prvního a posledního bodu v $[0, 1]$, kde $\beta(\cdot)$ kříží úroveň α

Blakerův algoritmus

- V Blaker (2000) krátký program v R, oprava Blaker (2001)

Blakerův algoritmus

- V Blaker (2000) krátký program v R, oprava Blaker (2001)
- Výpočet dolní konfidenční meze p_L:

Blakerův algoritmus

- V Blaker (2000) krátký program v R, oprava Blaker (2001)
- Výpočet dolní konfidenční meze p_L:
    1. Start v Clopper-Pearsonově mezi p_L^{CP}

Blakerův algoritmus

- V Blaker (2000) krátký program v R, oprava Blaker (2001)
- Výpočet dolní konfidenční meze p_L:
  1. Start v Clopper-Pearsonově mezi p_L^{CP}
  2. Iteruj $p := p + \Delta p$
     (konstantní krok $\Delta p > 0$, default 10^{-4}),
     dokud $\beta(p) < \alpha$

Blakerův algoritmus

- V Blaker (2000) krátký program v R, oprava Blaker (2001)
- Výpočet dolní konfidenční meze p_L:
  1. Start v Clopper-Pearsonově mezi p_L^{CP}
  2. Iteruj $p := p + \Delta p$
     (konstantní krok $\Delta p > 0$, default 10^{-4}),
     dokud $\beta(p) < \alpha$
  3. Když poprvé $\beta(p) \geq \alpha$, polož $p_L := p - \Delta p$, konec

Blakerův algoritmus

- V Blaker (2000) krátký program v R, oprava Blaker (2001)
- Výpočet dolní konfidenční meze p_L:
  1. Start v Clopper-Pearsonově mezi p_L^{CP}
  2. Iteruj $p := p + \Delta p$
     (konstantní krok $\Delta p > 0$, default 10^{-4}),
     dokud $\beta(p) < \alpha$
  3. Když poprvé $\beta(p) \geq \alpha$, polož $p_L := p - \Delta p$, konec
- Výpočet horní meze p_U analogický

Blakerův algoritmus

- Numerika je poněkud odbytá.

Blakerův algoritmus

- Numerika je poněkud odbytá.
- Konstantní krok $\rightarrow$ zvyšování přesnosti vede k drastickému zpomalení:

Blakerův algoritmus

- Numerika je poněkud odbytá.

- Konstantní krok $\rightarrow$ zvyšování přesnosti vede k drastickému zpomalení:

- Tabulky pro $n = 1, \ldots, 1000$ ($\sim$ 500 000 intervalů):

Blakerův algoritmus

- Numerika je poněkud odbytá.
- Konstantní krok $\rightarrow$ zvyšování přesnosti vede k drastickému zpomalení:
- Tabulky pro $n = 1, \ldots, 1000$ ($\sim 500\,000$ intervalů):
  - Tolerance 10^{-4} $\ldots \sim$ 6 min. na mém PC

Blakerův algoritmus

- Numerika je poněkud odbytá.
- Konstantní krok $\rightarrow$ zvyšování přesnosti vede k drastickému zpomalení:
- Tabulky pro $n = 1, \ldots, 1000$ ($\sim$ 500 000 intervalů):
  - Tolerance 10^{-4} $\ldots \sim$ 6 min. na mém PC
  - Tolerance 10^{-5} $\ldots \sim$ 53 min.

Blakerův algoritmus

- Numerika je poněkud odbytá.
- Konstantní krok $\rightarrow$ zvyšování přesnosti vede k drastickému zpomalení:
- Tabulky pro $n = 1, \ldots, 1000$ ($\sim 500\,000$ intervalů):
  - Tolerance 10^{-4} $\ldots \sim$ 6 min. na mém PC
  - Tolerance 10^{-5} $\ldots \sim$ 53 min.
  - Tolerance 10^{-6} $\ldots >$ 8 hod.

Blakerův algoritmus

- Numerika je poněkud odbytá.
- Konstantní krok $\rightarrow$ zvyšování přesnosti vede k drastickému zpomalení:
- Tabulky pro $n = 1, \ldots, 1000$ ($\sim 500\,000$ intervalů):
  - Tolerance 10^{-4} $\ldots \sim 6$ min. na mém PC
  - Tolerance 10^{-5} $\ldots \sim 53$ min.
  - Tolerance 10^{-6} $\ldots > 8$ hod.
  - Tolerance 10^{-7} $\ldots$ odhad $\sim$ půl týdne

Blakerův algoritmus

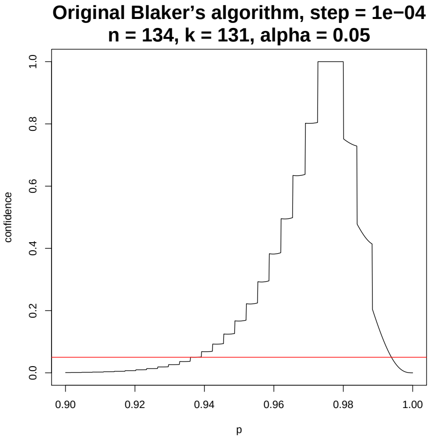
- Numerika je poněkud odbytá.
- Konstantní krok $\rightarrow$ zvyšování přesnosti vede k drastickému zpomalení:
- Tabulky pro $n = 1, \ldots, 1000$ ($\sim 500\,000$ intervalů):
  - Tolerance 10^{-4} $\ldots \sim 6$ min. na mém PC
  - Tolerance 10^{-5} $\ldots \sim 53$ min.
  - Tolerance 10^{-6} $\ldots > 8$ hod.
  - Tolerance 10^{-7} $\ldots$ odhad $\sim$ půl týdne
  - Tolerance 10^{-12} $\ldots$???

Blakerův algoritmus

- Numerika je poněkud odbytá.
- Konstantní krok $\rightarrow$ zvyšování přesnosti vede k drastickému zpomalení:
- Tabulky pro $n = 1, \ldots, 1000$ ($\sim 500\,000$ intervalů):
  - Tolerance 10^{-4} $\ldots \sim 6$ min. na mém PC
  - Tolerance 10^{-5} $\ldots \sim 53$ min.
  - Tolerance 10^{-6} $\ldots > 8$ hod.
  - Tolerance 10^{-7} $\ldots$odhad $\sim$ půl týdne
  - Tolerance 10^{-12} $\ldots$???
- Ještě hůř: Algoritmus je nepřesný – může „přehlédnout“ krátký interval, kde $\beta(p) > \alpha$, a skončit daleko od správného řešení (mnohem dále než o Δp).

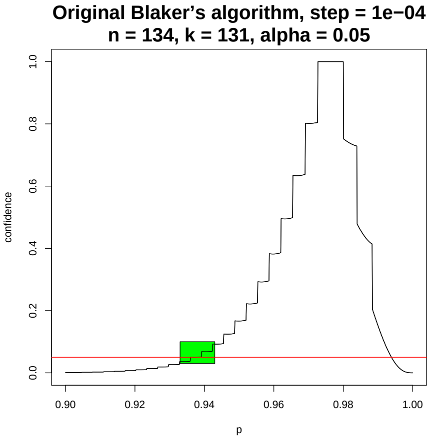
Blakerův algoritmus

- Příklad selhání:



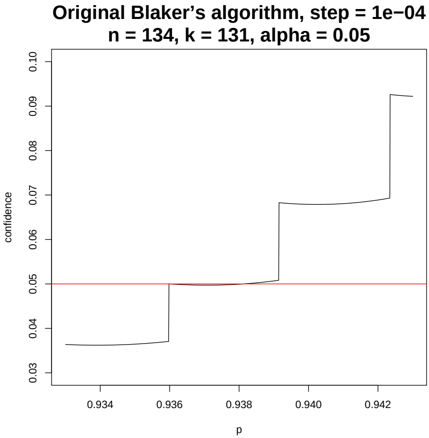
Blakerův algoritmus

- Příklad selhání:



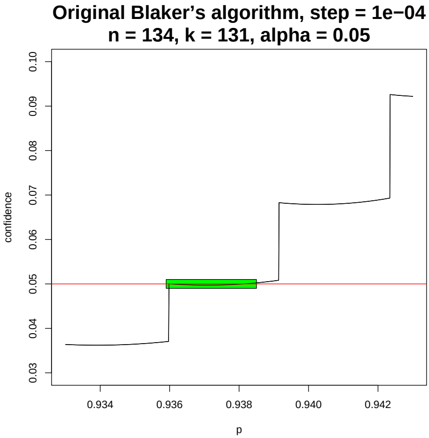
Blakerův algoritmus

- Příklad selhání:



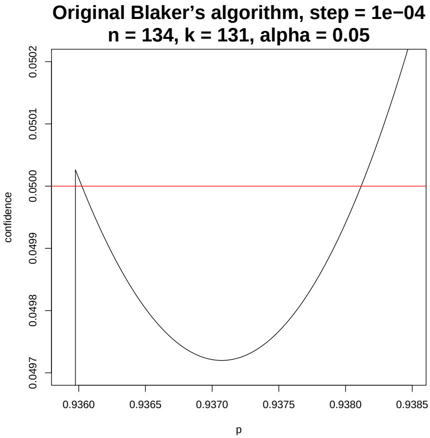
Blakerův algoritmus

- Příklad selhání:



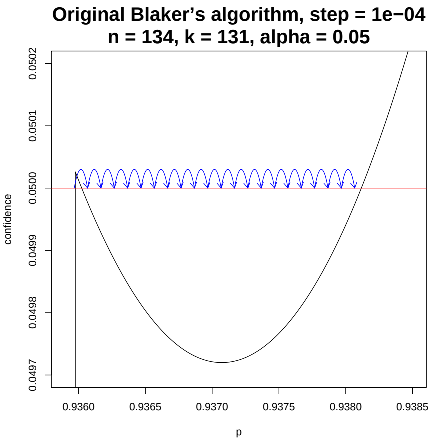
Blakerův algoritmus

- Příklad selhání:



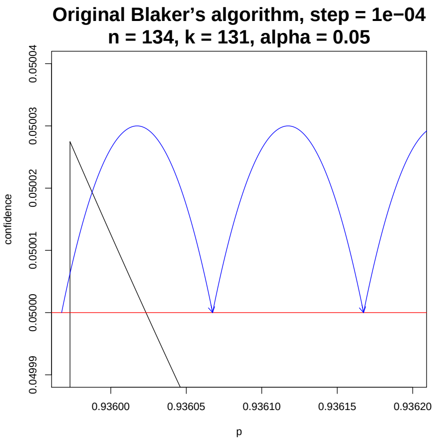
Blakerův algoritmus

- Příklad selhání:



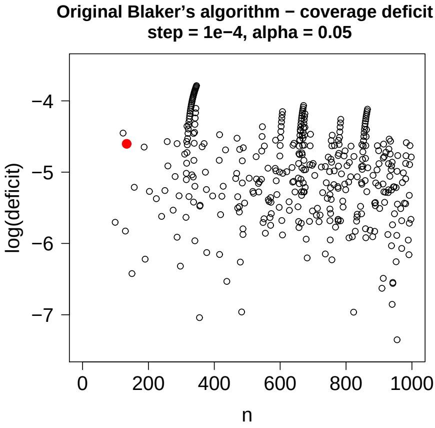
Blakerův algoritmus

- Příklad selhání:



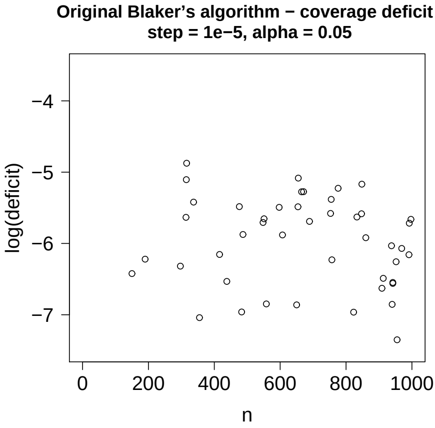
Blakerův algoritmus

- Statistika deficitů pokrytí – $n = 1, \ldots, 1000$:



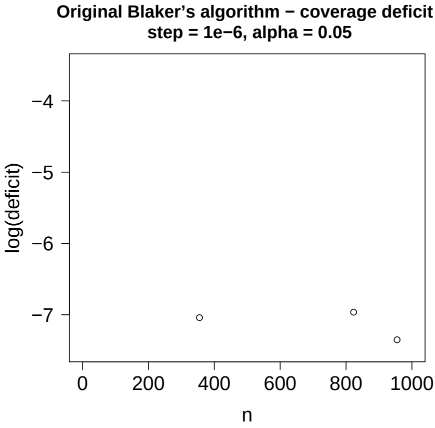
Blakerův algoritmus

- Statistika deficitů pokrytí – $n = 1, \ldots, 1000$:



Blakerův algoritmus

- Statistika deficitů pokrytí – $n = 1, \ldots, 1000$:



Nový algoritmus

- Přesněji řečeno, algoritmus byl nový na ROBUSTu 2010, teď už je nanejvýš zánovní.

Nový algoritmus

- Přesněji řečeno, algoritmus byl nový na ROBUSTu 2010, teď už je nanejvýš zánovní.
- Od ROBUSTu 2010 ale přeci jen doznal určitých změn.

Nový algoritmus

- Přesněji řečeno, algoritmus byl nový na ROBUSTu 2010, teď už je nanejvýš zánovní.
- Od ROBUSTu 2010 ale přeci jen doznal určitých změn.
- Může být výpočet Blakerových konfidenčních mezí rychlý, přesný a bez rizika selhání?

Nový algoritmus

- Přesněji řečeno, algoritmus byl nový na ROBUSTu 2010, teď už je nanejvýš zánovní.
- Od ROBUSTu 2010 ale přeci jen doznal určitých změn.
- Může být výpočet Blakerových konfidenčních mezí rychlý, přesný a bez rizika selhání?
  - Sotva, pokud s $\beta(\cdot)$ zacházíme jako s neznámou funkcí.

Nový algoritmus

- Přesněji řečeno, algoritmus byl nový na ROBUSTu 2010, teď už je nanejvýš zánovní.
- Od ROBUSTu 2010 ale přeci jen doznal určitých změn.
- Může být výpočet Blakerových konfidenčních mezí rychlý, přesný a bez rizika selhání?
  - Sotva, pokud s $\beta(\cdot)$ zacházíme jako s neznámou funkcí.
  - Ano, pokud dokážeme a využijeme některé vlastnosti $\beta(\cdot)$

Nový algoritmus – lemmata

- p_L^{CP} ... Clopper-Pearsonova dolní mez
  p^* ... první bod nespojitosti $\beta(\cdot)$ napravo od p_L^{CP}

Nový algoritmus – lemmata

- p_L^{CP} ... Clopper-Pearsonova dolní mez
  p^* ... první bod nespojitosti $\beta(\cdot)$ napravo od p_L^{CP}
- Lemma 1: $p_L^{CP} \leq p_L \leq p^*$

Nový algoritmus – lemmata

- p_L^{CP} ... Clopper-Pearsonova dolní mez
  p^* ... první bod nespojitosti $\beta(\cdot)$ napravo od p_L^{CP}
- Lemma 1: $p_L^{CP} \leq p_L \leq p^*$
- Lemma 2:
  Je-li $\beta(\cdot)$ spojitá na $[p_1, p_2]$ a $p_1 < p < p_2$,
  pak $\beta(p) < \max(\beta(p_1), \beta(p_2))$ (kvazikonvexita)

Nový algoritmus – lemmata

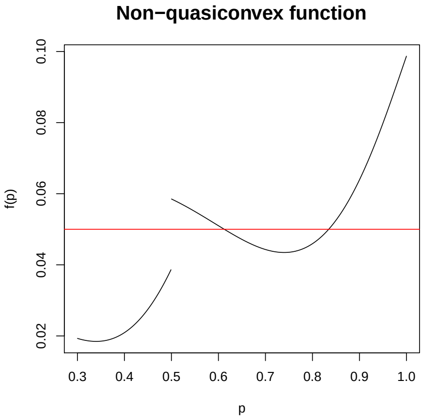
- p_L^{CP} ... Clopper-Pearsonova dolní mez
  p^* ... první bod nespojitosti $\beta(\cdot)$ napravo od p_L^{CP}
- Lemma 1: $p_L^{CP} \leq p_L \leq p^*$
- Lemma 2:
  Je-li $\beta(\cdot)$ spojitá na $[p_1, p_2]$ a $p_1 < p < p_2$,
  pak $\beta(p) < \max(\beta(p_1), \beta(p_2))$ (kvazikonvexita)
- Důsledky:
  - $\beta(\cdot)$ kříží úroveň α na $[p_L^{CP}, p^*]$ pouze jednou
  - Bod, kde se tak děje, lze bezpečně najít půlením intervalu

Nový algoritmus – lemmata

- p_L^{CP} ... Clopper-Pearsonova dolní mez
  p^* ... první bod nespojitosti $\beta(\cdot)$ napravo od p_L^{CP}
- Lemma 1: $p_L^{CP} \leq p_L \leq p^*$
- Lemma 2:
  Je-li $\beta(\cdot)$ spojitá na $[p_1, p_2]$ a $p_1 < p < p_2$,
  pak $\beta(p) < \max(\beta(p_1), \beta(p_2))$ (kvazikonvexita)
- Důsledky:
  - $\beta(\cdot)$ kříží úroveň α na $[p_L^{CP}, p^*]$ pouze jednou
  - Bod, kde se tak děje, lze bezpečně najít půlením intervalu
- Analogicky pro p_U

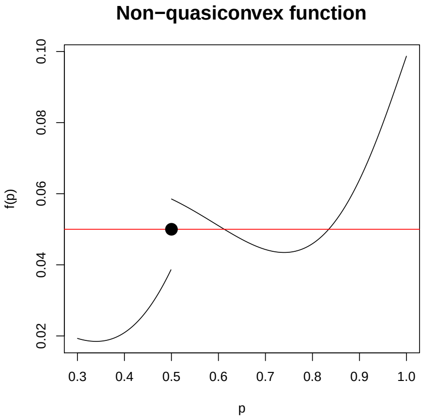
Nový algoritmus – kvazikonvexita a půlení

- Funkce, která není kvazikonvexní, může překročit hladinu α ve více bodech a při půlení intervalu může konvergovat k jinému z těchto bodů než k prvnímu zleva.



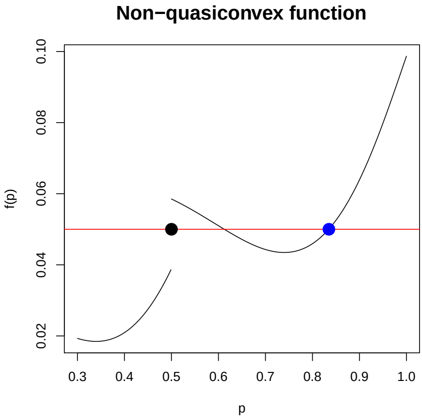
Nový algoritmus – kvazikonvexita a půlení

- Funkce, která není kvazikonvexní, může překročit hladinu α ve více bodech a při půlení intervalu může konvergovat k jinému z těchto bodů než k prvnímu zleva.



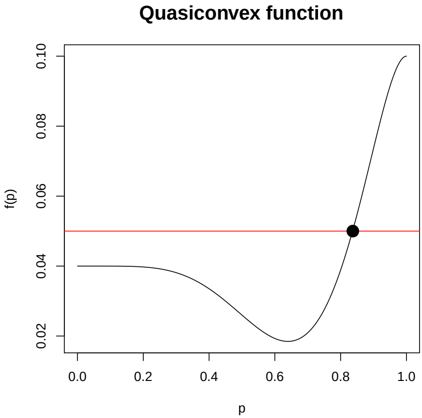
Nový algoritmus – kvazikonvexita a půlení

- Funkce, která není kvazikonvexní, může překročit hladinu α ve více bodech a při půlení intervalu může konvergovat k jinému z těchto bodů než k prvnímu zleva.



Nový algoritmus – kvazikonvexita a půlení

- Kvazikonvexní funkce nemůže po překročení hladiny α pod ni znovu klesnout, překročí ji proto v jediném bodě. Půlení intervalu tento bod bezpečně najde.



Nový algoritmus – popis

- Hledání dolní konfidenční meze p_L:
  (Pro p_U analogicky)

Nový algoritmus – popis

- Hledání dolní konfidenční meze p_L:
  (Pro p_U analogicky)

  1. Modifikuj $\beta(\cdot)$:

$$\beta^*(p) = \begin{cases} \beta(p) & p < p^* \\ 1 & p \geq p^* \end{cases}$$

Nový algoritmus – popis

- Hledání dolní konfidenční meze p_L:
  (Pro p_U analogicky)

1. Modifikuj $\beta(\cdot)$:

$$\beta^*(p) = \begin{cases} \beta(p) & p < p^* \\ 1 & p \geq p^* \end{cases}$$

2. Aplikuj na $\beta^*(\cdot)$ mezi p_L^{CP} a 1 půlení intervalu

Nový algoritmus – popis

- Hledání dolní konfidenční meze p_L:
  (Pro p_U analogicky)

  1. Modifikuj $\beta(\cdot)$:

  $$\beta^*(p) = \begin{cases} \beta(p) & p < p^* \\ 1 & p \geq p^* \end{cases}$$

  2. Aplikuj na $\beta^*(\cdot)$ mezi p_L^{CP} a 1 půlení intervalu

- Test, zda $p \geq p^*$, je snadný: $P_p(X \geq k) \geq P_p(X \leq k_p^* + 1)$?

Nový algoritmus – popis

- Hledání dolní konfidenční meze p_L:
  (Pro p_U analogicky)

  1. Modifikuj $\beta(\cdot)$:

  $$\beta^*(p) = \begin{cases} \beta(p) & p < p^* \\ 1 & p \geq p^* \end{cases}$$

  2. Aplikuj na $\beta^*(\cdot)$ mezi p_L^{CP} a 1 půlení intervalu

- Test, zda $p \geq p^*$, je snadný: $P_p(X \geq k) \geq P_p(X \leq k_p^* + 1)$?

- $\beta^*(\cdot)$ je kvazikonvexní na $[p_L^{CP}, 1] \longrightarrow$ půlení intervalu
  je bezpečné „globálně“, nejen na $[p_L^{CP}, p^*]$.

Nový algoritmus – závěrečné poznámky

- Algoritmus je rychlý a přesný

Nový algoritmus – závěrečné poznámky

- Algoritmus je rychlý a přesný
- Tabulky pro $n = 1, \ldots, 1000$ ($\sim 500\,000$ intervalů):

Nový algoritmus – závěrečné poznámky

- Algoritmus je rychlý a přesný
- Tabulky pro $n = 1, \dots, 1000$ ($\sim 500\,000$ intervalů):
  - Tolerance 10^{-4} $\dots \sim 6$ min. na mém PC

Nový algoritmus – závěrečné poznámky

- Algoritmus je rychlý a přesný
- Tabulky pro $n = 1, \dots, 1000$ ($\sim 500\,000$ intervalů):
  - Tolerance 10^{-4} $\dots \sim$ 6 min. na mém PC
  - ...

Nový algoritmus – závěrečné poznámky

- Algoritmus je rychlý a přesný
- Tabulky pro $n = 1, \dots, 1000$ ($\sim 500\,000$ intervalů):
  - Tolerance 10^{-4} $\dots \sim$ 6 min. na mém PC
  - $\dots$
  - Tolerance 10^{-12} $\dots \sim$ 23 min.

Nový algoritmus – závěrečné poznámky

- Algoritmus je rychlý a přesný
- Tabulky pro $n = 1, \ldots, 1000$ ($\sim 500\,000$ intervalů):
  - Tolerance 10^{-4} $\ldots \sim$ 6 min. na mém PC
  - $\ldots$
  - Tolerance 10^{-12} $\ldots \sim$ 23 min.
  - Bez deficitů pokrytí

Nový algoritmus – závěrečné poznámky

- Algoritmus je rychlý a přesný
- Tabulky pro $n = 1, \ldots, 1000$ ($\sim 500\,000$ intervalů):
  - Tolerance 10^{-4} $\ldots \sim 6$ min. na mém PC
  - $\ldots$
  - Tolerance 10^{-12} $\ldots \sim 23$ min.
  - Bez deficitů pokrytí
- Implementace: Několik desítek řádek v R – v době ROBUSTu 2010 pracovní verze, později balíček `BlakerCI`.

Nový algoritmus – závěrečné poznámky

- Algoritmus je rychlý a přesný
- Tabulky pro $n = 1, \ldots, 1000$ ($\sim 500\,000$ intervalů):
  - Tolerance 10^{-4} $\ldots \sim$ 6 min. na mém PC
  - $\ldots$
  - Tolerance 10^{-12} $\ldots \sim$ 23 min.
  - Bez deficitů pokrytí
- Implementace: Několik desítek řádek v R – v době ROBUSTu 2010 pracovní verze, později balíček `BlakerCI`.
- To ale už do opakovací lekce nepatří.

Balíček `BlakerCI`

- Balíček `BlakerCI` je vystaven na CRANu (`http://cran.r-project.org/`).

Balíček `BlakerCI`

- Balíček `BlakerCI` je vystaven na CRANu (`http://cran.r-project.org/`).
  - 1. verze od listopadu 2010,

Balíček `BlakerCI`

- Balíček `BlakerCI` je vystaven na CRANu (`http://cran.r-project.org/`).
  - 1. verze od listopadu 2010,
  - 2. verze (jen nepatrně odlišná) od ledna 2011,

Balíček `BlakerCI`

- Balíček `BlakerCI` je vystaven na CRANu (`http://cran.r-project.org/`).
  - 1. verze od listopadu 2010,
  - 2. verze (jen nepatrně odlišná) od ledna 2011,
  - 3. verze (podstatně upravená) se chystá a chystá…

Balíček `BlakerCI`

- Balíček `BlakerCI` je vystaven na CRANu (`http://cran.r-project.org/`).
  - 1. verze od listopadu 2010,
  - 2. verze (jen nepatrně odlišná) od ledna 2011,
  - 3. verze (podstatně upravená) se chystá a chystá...

- Balíček je malý – kromě Blakerových binomických konfidenčních mezí dnes „umí" jen rychlý výpočet nejlepší horní aproximace funkce $\beta(\cdot)$ unimodální funkcí. (Použití: p-hodnota testu, který invertuje Blakerův interval.)

Balíček `BlakerCI`

- Balíček `BlakerCI` je vystaven na CRANu (`http://cran.r-project.org/`).
  - 1. verze od listopadu 2010,
  - 2. verze (jen nepatrně odlišná) od ledna 2011,
  - 3. verze (podstatně upravená) se chystá a chystá...

- Balíček je malý – kromě Blakerových binomických konfidenčních mezí dnes „umí“ jen rychlý výpočet nejlepší horní aproximace funkce $\beta(\cdot)$ unimodální funkcí. (Použití: p-hodnota testu, který invertuje Blakerův interval.)

- V plánu rozšiřování o další rozdělení/úlohy.

Balíček `BlakerCI`

- Balíček `BlakerCI` je vystaven na CRANu (`http://cran.r-project.org/`).
  - 1. verze od listopadu 2010,
  - 2. verze (jen nepatrně odlišná) od ledna 2011,
  - 3. verze (podstatně upravená) se chystá a chystá...

- Balíček je malý – kromě Blakerových binomických konfidenčních mezí dnes „umí“ jen rychlý výpočet nejlepší horní aproximace funkce $\beta(\cdot)$ unimodální funkcí. (Použití: p-hodnota testu, který invertuje Blakerův interval.)

- V plánu rozšiřování o další rozdělení/úlohy.

- Váhám, jestli zařadit další pomocný materiál, např.

Balíček `BlakerCI`

- Balíček `BlakerCI` je vystaven na CRANu (`http://cran.r-project.org/`).
  - 1. verze od listopadu 2010,
  - 2. verze (jen nepatrně odlišná) od ledna 2011,
  - 3. verze (podstatně upravená) se chystá a chystá...
- Balíček je malý – kromě Blakerových binomických konfidenčních mezí dnes „umí“ jen rychlý výpočet nejlepší horní aproximace funkce $\beta(\cdot)$ unimodální funkcí. (Použití: p-hodnota testu, který invertuje Blakerův interval.)
- V plánu rozšiřování o další rozdělení/úlohy.
- Váhám, jestli zařadit další pomocný materiál, např.
  - programy pro testování vypočtených CI,

Balíček BlakerCI

- Balíček `BlakerCI` je vystaven na CRANu (`http://cran.r-project.org/`).
  - 1. verze od listopadu 2010,
  - 2. verze (jen nepatrně odlišná) od ledna 2011,
  - 3. verze (podstatně upravená) se chystá a chystá…
- Balíček je malý – kromě Blakerových binomických konfidenčních mezí dnes „umí“ jen rychlý výpočet nejlepší horní aproximace funkce $\beta(\cdot)$ unimodální funkcí. (Použití: p-hodnota testu, který invertuje Blakerův interval.)
- V plánu rozšiřování o další rozdělení/úlohy.
- Váhám, jestli zařadit další pomocný materiál, např.
  - programy pro testování vypočtených CI,
  - programy pro výpočet tabulek CI, popř. i tyto tabulky.

Práce M. P. Faye

- *Confidence intervals that match Fisher's exact or Blaker's exact tests*. Biostatistics (2010)

Práce M. P. Faye

- *Confidence intervals that match Fisher's exact or Blaker's exact tests*. Biostatistics (2010)
- *Two-sided exact tests and matching confidence intervals for discrete data*. R journal (2010)

Práce M. P. Faye

- *Confidence intervals that match Fisher's exact or Blaker's exact tests*. Biostatistics (2010)
- *Two-sided exact tests and matching confidence intervals for discrete data*. R journal (2010)
- Studuje několik jedno- a dvouvýběrových úloh pro diskrétní data

Práce M. P. Faye

- *Confidence intervals that match Fisher's exact or Blaker's exact tests*. Biostatistics (2010)
- *Two-sided exact tests and matching confidence intervals for discrete data*. R journal (2010)
- Studuje několik jedno- a dvouvýběrových úloh pro diskrétní data
- Rozeznává 3 typy oboustranných testů a jim odpovídajících konfidenčních intervalů – tzv. `central`, `minlike` a `blaker`

Práce M. P. Faye

- *Confidence intervals that match Fisher's exact or Blaker's exact tests*. Biostatistics (2010)
- *Two-sided exact tests and matching confidence intervals for discrete data*. R journal (2010)
- Studuje několik jedno- a dvouvýběrových úloh pro diskrétní data
- Rozeznává 3 typy oboustranných testů a jim odpovídajících konfidenčních intervalů – tzv. `central`, `minlike` a `blaker`
- Kritizuje, že standardní SW (SAS, R) užívá testy a CI nestejného typu, odtud časté inkonsistentní výsledky – zamítnutá hodnota parametru je v CI, popř. naopak

Práce M. P. Faye

- Nabízí balíky `exact2x2` a `exactci` v R –
  spolu s testem se vždy počítá CI odpovídajícího typu

Práce M. P. Faye

- Nabízí balíky `exact2x2` a `exactci` v R –
  spolu s testem se vždy počítá CI odpovídajícího typu
- Balík `exactci` počítá mj. i Blakerův CI pro parametr binomického rozdělení

Práce M. P. Faye

- Nabízí balíky `exact2x2` a `exactci` v R –
  spolu s testem se vždy počítá CI odpovídajícího typu
- Balík `exactci` počítá mj. i Blakerův CI pro parametr binomického rozdělení
- Fay upozorňuje na nepřesnost Blakerova algoritmu, navrhuje svůj vlastní

Algoritmus M. P. Faye

- Hledá dostatečně krátký interval, kde levá (pravá) konfidenční mez p_L (p_U) musí ležet.

Algoritmus M. P. Faye

- Hledá dostatečně krátký interval, kde levá (pravá) konfidenční mez p_L (p_U) musí ležet.
- Začíná dělením $[0, 1]$ pomocí bodů nespojitosti funkce $\beta(\cdot)$. „Zajímavé“ intervaly (kde hledaný bod může nebo musí být) dělí dále jemněji.

Algoritmus M. P. Faye

- Hledá dostatečně krátký interval, kde levá (pravá) konfidenční mez p_L (p_U) musí ležet.
- Začíná dělením $[0, 1]$ pomocí bodů nespojitosti funkce $\beta(\cdot)$. „Zajímavé" intervaly (kde hledaný bod může nebo musí být) dělí dále jemněji.
- O eliminaci nebo podrobnějším zkoumání intervalu I se rozhoduje podle dolního odhadu minima a horního odhadu maxima funkce $\beta(\cdot)$ na I.

Algoritmus M. P. Faye

- Hledá dostatečně krátký interval, kde levá (pravá) konfidenční mez p_L (p_U) musí ležet.
- Začíná dělením $[0, 1]$ pomocí bodů nespojitosti funkce $\beta(\cdot)$. „Zajímavé“ intervaly (kde hledaný bod může nebo musí být) dělí dále jemněji.
- O eliminaci nebo podrobnějším zkoumání intervalu I se rozhoduje podle dolního odhadu minima a horního odhadu maxima funkce $\beta(\cdot)$ na I.
- Odhady jsou založeny na tom, že $\beta(\cdot)$ je v intervalu, kde je spojitá, součtem dvou známých funkcí, z nichž jedna je rostoucí a druhá klesající.

Přednosti a nedostatky algoritmu M. P. Faye

- Předností je obecnost – funkce `exactbinomCI` počítá obdobným způsobem dva různé typy CI pro parametr binomického rozdělení a má společnou „kostru“ s funkcemi určenými pro jiné úlohy / jiná rozdělení.

Přednosti a nedostatky algoritmu M. P. Faye

- Předností je obecnost – funkce `exactbinomCI` počítá obdobným způsobem dva různé typy CI pro parametr binomického rozdělení a má společnou „kostru“ s funkcemi určenými pro jiné úlohy / jiná rozdělení.

- Fay nevyužívá některé užitečné vlastnosti funkce $\beta(\cdot)$
  $\longrightarrow$ jeho algoritmus v dané úloze dělá spoustu zbytečné práce
  $\longrightarrow$ je pomalý, zvláště pro velká n.

Přednosti a nedostatky algoritmu M. P. Faye

- Předností je obecnost – funkce `exactbinomCI` počítá obdobným způsobem dva různé typy CI pro parametr binomického rozdělení a má společnou „kostru“ s funkcemi určenými pro jiné úlohy / jiná rozdělení.

- Fay nevyužívá některé užitečné vlastnosti funkce $\beta(\cdot)$
$\longrightarrow$ jeho algoritmus v dané úloze dělá spoustu zbytečné práce
$\longrightarrow$ je pomalý, zvláště pro velká n.

- Tabulky 95% CI pro $n = 1, \ldots, 1000$ ($\approx 500\,000$ intervalů):
Na clusteru $\sim$ 48 hod. (můj program $\sim$ 10 min. – cca $300 : 1$).

Přednosti a nedostatky algoritmu M. P. Faye

- Předností je obecnost – funkce `exactbinomCI` počítá obdobným způsobem dva různé typy CI pro parametr binomického rozdělení a má společnou „kostru“ s funkcemi určenými pro jiné úlohy / jiná rozdělení.

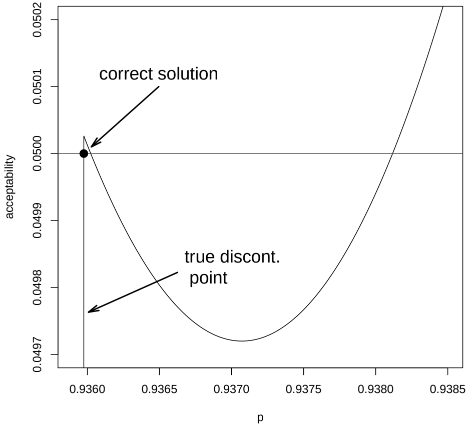
- Fay nevyužívá některé užitečné vlastnosti funkce $\beta(\cdot)$
$\longrightarrow$ jeho algoritmus v dané úloze dělá spoustu zbytečné práce
$\longrightarrow$ je pomalý, zvláště pro velká n.

- Tabulky 95% CI pro $n = 1, \ldots, 1000$ ($\approx 500\,000$ intervalů): Na clusteru $\sim$ 48 hod. (můj program $\sim$ 10 min. – cca 300 : 1).

- Fayův algoritmus je přesnější než Blakerův, ale také není imunní vůči větším chybám plynoucím z toho, že „přehlédne“ (popř. naopak „vyčaruje“) krátký interval, kde $\beta(p) > \alpha$.

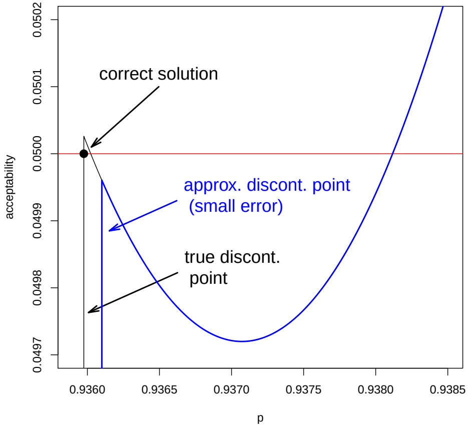
Přednosti a nedostatky algoritmu M. P. Faye

- Numericky s danou přesností se počítají body nespojitosti $\beta(\cdot)$ $\longrightarrow$ lze „přehlédnout“ krátký interval, kde $\beta(p) > \alpha$:



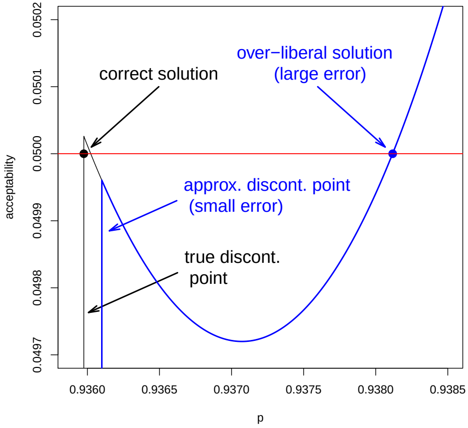
Přednosti a nedostatky algoritmu M. P. Faye

- Numericky s danou přesností se počítají body nespojitosti $\beta(\cdot)$ $\longrightarrow$ lze „přehlédnout“ krátký interval, kde $\beta(p) > \alpha$:



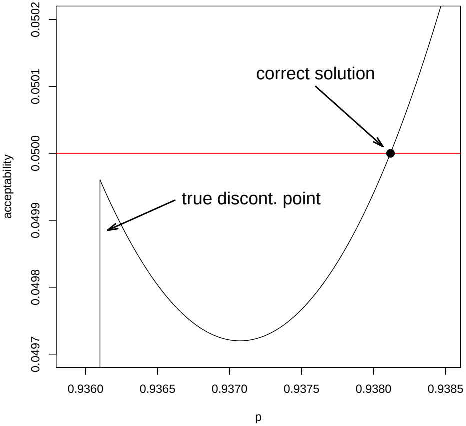
Přednosti a nedostatky algoritmu M. P. Faye

- Numericky s danou přesností se počítají body nespojitosti $\beta(\cdot)$ $\longrightarrow$ lze „přehlédnout“ krátký interval, kde $\beta(p) > \alpha$:



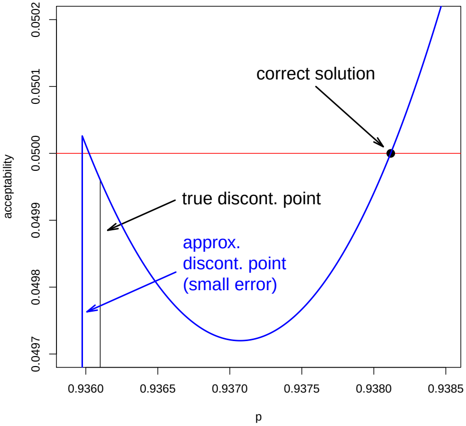
Přednosti a nedostatky algoritmu M. P. Faye

- Numerickým výpočtem bodů nespojitosti $\beta(\cdot)$
  lze krátký interval, kde $\beta(p) > \alpha$, i „přičarovat“:



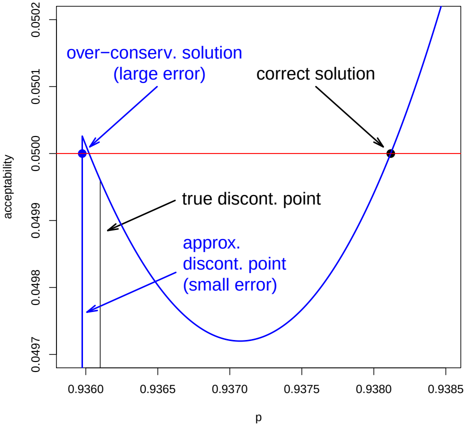
Přednosti a nedostatky algoritmu M. P. Faye

- Numerickým výpočtem bodů nespojitosti $\beta(\cdot)$ lze krátký interval, kde $\beta(p) > \alpha$, i „přičarovat“:



Přednosti a nedostatky algoritmu M. P. Faye

- Numerickým výpočtem bodů nespojitosti $\beta(\cdot)$ lze krátký interval, kde $\beta(p) > \alpha$, i „přičarovat“:



Přednosti a nedostatky algoritmu M. P. Faye

- Fayův algoritmus počítá body nespojitosti funkce $\beta(\cdot)$ numericky (funkcí `uniroot`) s danou numerickou tolerancí.

Přednosti a nedostatky algoritmu M. P. Faye

- Fayův algoritmus počítá body nespojitosti funkce $\beta(\cdot)$ numericky (funkcí `uniroot`) s danou numerickou tolerancí.
- Nepřesnosti předvedené v grafech jsou daleko vzácněji než u Blakerova algoritmu, ale reálně se vyskytují.

Přednosti a nedostatky algoritmu M. P. Faye

- Fayův algoritmus počítá body nespojitosti funkce $\beta(\cdot)$ numericky (funkcí `uniroot`) s danou numerickou tolerancí.
- Nepřesnosti předvedené v grafech jsou daleko vzácněji než u Blakerova algoritmu, ale reálně se vyskytují.
- Můj algoritmus body nespojitosti p^* explicitně nepočítá, testuje jen $p \geq p^*$ (nepřímo – pomocí testu, zda $P_p(X \geq k) \geq P_p(X \leq k_p^* + 1)$).

Radosti počítání s konečnou přesností

- Tvrzení, že můj algoritmus je přesný, je třeba upřesnit.

Radosti počítání s konečnou přesností

- Tvrzení, že můj algoritmus je přesný, je třeba upřesnit.
- Samozřejmě počítá konfidenční meze se zadanou numerickou tolerancí. Ale i to s určitým rizikem.

Radosti počítání s konečnou přesností

- Tvrzení, že můj algoritmus je přesný, je třeba upřesnit.
- Samozřejmě počítá konfidenční meze se zadanou numerickou tolerancí. Ale i to s určitým rizikem.
- Riziko by nebylo, kdyby (např.)

Radosti počítání s konečnou přesností

- Tvrzení, že můj algoritmus je přesný, je třeba upřesnit.
- Samozřejmě počítá konfidenční meze se zadanou numerickou tolerancí. Ale i to s určitým rizikem.
- Riziko by nebylo, kdyby (např.)
  - byl absolutně přesný výpočet distribuční funkce binomického rozdělení;

Radosti počítání s konečnou přesností

- Tvrzení, že můj algoritmus je přesný, je třeba upřesnit.
- Samozřejmě počítá konfidenční meze se zadanou numerickou tolerancí. Ale i to s určitým rizikem.
- Riziko by nebylo, kdyby (např.)
  - byl absolutně přesný výpočet distribuční funkce binomického rozdělení;
  - se absolutně přesně počítaly Clopper-Pearsonovy konfidenční meze (tj. kvantily rozdělení beta);

Radosti počítání s konečnou přesností

- Tvrzení, že můj algoritmus je přesný, je třeba upřesnit.
- Samozřejmě počítá konfidenční meze se zadanou numerickou tolerancí. Ale i to s určitým rizikem.
- Riziko by nebylo, kdyby (např.)
  - byl absolutně přesný výpočet distribuční funkce binomického rozdělení;
  - se absolutně přesně počítaly Clopper-Pearsonovy konfidenční meze (tj. kvantily rozdělení beta);
  - se nemohlo stát, že Clopper-Pearsonovu mez přiřadíme k jiné spojité části funkce $\beta(\cdot)$, než kam patří ($\sim$ kvantily binomického rozdělení).

Přednosti a nedostatky algoritmu M. P. Faye

- Vadí ale i jednodušší věci:

Přednosti a nedostatky algoritmu M. P. Faye

- Vadí ale i jednodušší věci:
  - Když zadáme hladinu spolehlivosti 0.95, má se zkoumat, kde $\beta(p) > 0.05$.

Přednosti a nedostatky algoritmu M. P. Faye

- Vadí ale i jednodušší věci:
  - Když zadáme hladinu spolehlivosti 0.95, má se zkoumat, kde $\beta(p) > 0.05$.
  - Jenže $1 - 0.05$ není (v aritmetice užívané v R) totéž co 0.05, nýbrž o cca 4×10^{-17} víc.

Přednosti a nedostatky algoritmu M. P. Faye

- Vadí ale i jednodušší věci:
  - Když zadáme hladinu spolehlivosti 0.95, má se zkoumat, kde $\beta(p) > 0.05$.
  - Jenže $1 - 0.05$ není (v aritmetice užívané v R) totéž co 0.05, nýbrž o cca 4×10^{-17} víc.
  - Závislost konfidenčních mezí p_L, p_U na hladině spolehlivosti $1 - \alpha$ není spojitá
    $\longrightarrow$ „mikroskopická" chyba v α může způsobit „makroskopickou" chybu v p_L, p_U.

Přednosti a nedostatky algoritmu M. P. Faye

- Vadí ale i jednodušší věci:
  - Když zadáme hladinu spolehlivosti 0.95, má se zkoumat, kde $\beta(p) > 0.05$.
  - Jenže $1 - 0.05$ není (v aritmetice užívané v R) totéž co 0.05, nýbrž o cca 4×10^{-17} víc.
  - Závislost konfidenčních mezí p_L, p_U na hladině spolehlivosti $1 - \alpha$ není spojitá $\longrightarrow$ „mikroskopická" chyba v α může způsobit „makroskopickou" chybu v p_L, p_U.
- Ve stávající verzi balíčku `BlakerCI` se pravá konfidenční mez $p_U(k)$ pro $X = k$ počítá na základě symetrie jako $1 - p_L(n - k)$. Už jen to způsobuje, že testy „správnosti" mezí hlásí občas drobné problémy.

Děkuji za pozornost