

Technika znáhodněných odpovědí

Jiří Dvořák

Katedra pravděpodobnosti a matematické statistiky

Soustředění mladých fyziků a matematiků

1. července 2025

Jak zjistit podíl lidí v této místnosti, kteří nosí brýle?



Jak zjistit podíl lidí v této místnosti, kteří mají zvířátko?



Jaký byl rozdíl mezi těmi problémy?



Jak zjistit podíl lidí v této místnosti, kteří dali úplatek?



Jaký byl rozdíl mezi těmi problémy?



Odbočka, co vlastně není odbočka

Jak zjistit podíl dospělých v ČR, kteří podporují politickou stranu XYZ?

Odbočka, co vlastně není odbočka

Jak zjistit podíl dospělých v ČR, kteří podporují politickou stranu XYZ?

Výběr vzorku, pak odhad

$$\hat{p} = \frac{ANO}{ANO + NE}$$

Odbočka, co vlastně není odbočka

Jak zjistit podíl dospělých v ČR, kteří podporují politickou stranu XYZ?

Výběr vzorku, pak odhad

$$\hat{p} = \frac{ANO}{ANO + NE}$$

Pozor: dobrý způsob výběru vzorku je napůl věda, napůl umění.

- ananas na pizze,

- ananas na pizze,
- nezákonné chování,
- mentální zdraví,
- sexuální chování,
- osobní hygiena,
- ...

- ananas na pizze,
- nezákonné chování,
- mentální zdraví,
- sexuální chování,
- osobní hygiena,
- ...

Snaha zajistit, aby tazatel nezjistil pravdu o respondentovi.

Zrcadlová otázka (1965)

Respondent bude odpovídat na citlivou otázku (A) či její opak (B).

Zrcadlová otázka (1965)

Respondent bude odpovídat na citlivou otázku (A) či její opak (B).

Rozhodne o tom náhoda (hod kostkou, zatočení šipkou, ...).

Zrcadlová otázka (1965)

Respondent bude odpovídat na citlivou otázku (A) či její opak (B).

Rozhodne o tom náhoda (hod kostkou, zatočení šipkou, . . .).

Tazatel nesmí vědět, na kterou otázku respondent odpovídá!

Korupce (Latinská Amerika, 2010, cca 3000 respondentů,
30 institucí):

For each of the following questions, please spin the arrow until it
has made at least one full rotation.

Korupce (Latinská Amerika, 2010, cca 3000 respondentů,
30 institucí):

For each of the following questions, please spin the arrow until it has made at least one full rotation.

If the arrow lands on region A for a particular question, respond true or false in the space indicated only with respect to statement A.

Korupce (Latinská Amerika, 2010, cca 3000 respondentů,
30 institucí):

For each of the following questions, please spin the arrow until it has made at least one full rotation.

If the arrow lands on region A for a particular question, respond true or false in the space indicated only with respect to statement A.

If the arrow lands on region B for that question, respond true or false in the space indicated only with respect to statement B.

Korupce (Latinská Amerika, 2010, cca 3000 respondentů, 30 institucí):

For each of the following questions, please spin the arrow until it has made at least one full rotation.

If the arrow lands on region A for a particular question, respond true or false in the space indicated only with respect to statement A.

If the arrow lands on region B for that question, respond true or false in the space indicated only with respect to statement B.

Do not make any marks to indicate in which region the arrow fell for each question. Please remember that if you respond false to a statement in its negative form that means that the positive form of the statement is true.

A. I have never used, not even once, the resources of my institution for the benefit of a political party.

B. I have used, at least once, the resources of my institution for the benefit of a political party.

Zrcadlová otázka – vyhodnocení

Chceme zjistit (odhadnout) hodnotu a = pravděpodobnost, že správná odpověď respondenta i na otázku A je „ano“.

Zrcadlová otázka – vyhodnocení

Chceme zjistit (odhadnout) hodnotu a = pravděpodobnost, že správná odpověď respondenta i na otázku A je „ano“.

p_A = pravděpodobnost, že respondent odpovídá na otázku A.

Zrcadlová otázka – vyhodnocení

Chceme zjistit (odhadnout) hodnotu a = pravděpodobnost, že správná odpověď respondenta i na otázku A je „ano“.

p_A = pravděpodobnost, že respondent odpovídá na otázku A.

p_{ANO} = pravděpodobnost, že od respondenta i uslyšíme „ano“.

Zrcadlová otázka – vyhodnocení

Chceme zjistit (odhadnout) hodnotu a = pravděpodobnost, že správná odpověď respondenta i na otázku A je „ano“.

p_A = pravděpodobnost, že respondent odpovídá na otázku A.

p_{ANO} = pravděpodobnost, že od respondenta i uslyšíme „ano“.

$$p_{ANO} = p_A \cdot a + (1 - p_A) \cdot (1 - a)$$

Zrcadlová otázka – vyhodnocení

Chceme zjistit (odhadnout) hodnotu a = pravděpodobnost, že správná odpověď respondenta i na otázku A je „ano“.

p_A = pravděpodobnost, že respondent odpovídá na otázku A.

p_{ANO} = pravděpodobnost, že od respondenta i uslyšíme „ano“.

$$p_{ANO} = p_A \cdot a + (1 - p_A) \cdot (1 - a)$$

$$a = \frac{1}{2p_A - 1} (p_{ANO} + p_A - 1)$$

Zrcadlová otázka – vyhodnocení

Chceme zjistit (odhadnout) hodnotu a = pravděpodobnost, že správná odpověď respondenta i na otázku A je „ano“.

p_A = pravděpodobnost, že respondent odpovídá na otázku A.

p_{ANO} = pravděpodobnost, že od respondenta i uslyšíme „ano“.

$$p_{ANO} = p_A \cdot a + (1 - p_A) \cdot (1 - a)$$

$$a = \frac{1}{2p_A - 1} (p_{ANO} + p_A - 1)$$

Pozor na $p_A = \frac{1}{2}$ (hod mincí), pak $p_{ANO} = \frac{1}{2}$ a nezjistíme nic!

Zrcadlová otázka – pizza

Hod'te kostkou. Pokud padne číslo 1, 2, 3, 4 nebo 5, vyjádřete svůj souhlas či nesouhlas s výrokem A. Pokud padne číslo 6, vyjádřete svůj souhlas či nesouhlas s výrokem B.

A. Ananas na pizzu patří.

B. Ananas na pizzu nepatří.

Zrcadlová otázka – pizza

Hod'te kostkou. Pokud padne číslo 1, 2, 3, 4 nebo 5, vyjádřete svůj souhlas či nesouhlas s výrokem A. Pokud padne číslo 6, vyjádřete svůj souhlas či nesouhlas s výrokem B.

A. Ananas na pizzu patří.

B. Ananas na pizzu nepatří.

$$\hat{a} = \frac{3}{2} \left(\hat{p}_{ANO} - \frac{1}{6} \right)$$

Vynucená odpověď (1971)

Náhoda rozhoduje o tom, zda respondent odpovídá na citlivou otázku nebo sděluje zadanou (vynucenou) odpověď.

Vynucená odpověď (1971)

Náhoda rozhoduje o tom, zda respondent odpovídá na citlivou otázku nebo sděluje zadanou (vynucenou) odpověď.

Pro respondenta méně matoucí než zrcadlová otázka.

Kontakty s militantními skupinami (Nigérie, 2014):

For this question, I want you to answer yes or no. But I want you to consider the number of your dice throw.

Kontakty s militantními skupinami (Nigérie, 2014):

For this question, I want you to answer yes or no. But I want you to consider the number of your dice throw.

If 1 shows on the dice, tell me no. If 6 shows, tell me yes. But if another number, like 2 or 3 or 4 or 5 shows, tell me your own opinion about the question that I will ask you after you throw the dice.

Kontakty s militantními skupinami (Nigérie, 2014):

For this question, I want you to answer yes or no. But I want you to consider the number of your dice throw.

If 1 shows on the dice, tell me no. If 6 shows, tell me yes. But if another number, like 2 or 3 or 4 or 5 shows, tell me your own opinion about the question that I will ask you after you throw the dice.

[TURN AWAY FROM THE RESPONDENT] Now you throw the dice so that I cannot see what comes out. Please do not forget the number that comes out. [WAIT TO TURN AROUND UNTIL RESPONDENT SAYS YES TO:] Have you thrown the dice? Have you picked it up?

Now, during the height of the conflict in 2007 and 2008, did you know any militants, like a family member, a friend, or someone you talked to on a regular basis. Please, before you answer, take note of the number you rolled on the dice.

Vynucená odpověď – vyhodnocení

Chceme zjistit (odhadnout) hodnotu a = pravděpodobnost, že správná odpověď respondenta i na citlivou otázku je „ano“.

Vynucená odpověď – vyhodnocení

Chceme zjistit (odhadnout) hodnotu a = pravděpodobnost, že správná odpověď respondenta i na citlivou otázku je „ano“.

$p_{V,ano}$ = pravděpodobnost, že uslyšíme vynucenou odpověď „ano“.

Vynucená odpověď – vyhodnocení

Chceme zjistit (odhadnout) hodnotu a = pravděpodobnost, že správná odpověď respondenta i na citlivou otázku je „ano“.

$p_{V,ano}$ = pravděpodobnost, že uslyšíme vynucenou odpověď „ano“.

$p_{V,ne}$ = pravděpodobnost, že uslyšíme vynucenou odpověď „ne“.

Vynucená odpověď – vyhodnocení

Chceme zjistit (odhadnout) hodnotu a = pravděpodobnost, že správná odpověď respondenta i na citlivou otázku je „ano“.

$p_{V,ano}$ = pravděpodobnost, že uslyšíme vynucenou odpověď „ano“.

$p_{V,ne}$ = pravděpodobnost, že uslyšíme vynucenou odpověď „ne“.

p_{ANO} = pravděpodobnost, že od respondenta i uslyšíme „ano“.

Vynucená odpověď – vyhodnocení

Chceme zjistit (odhadnout) hodnotu a = pravděpodobnost, že správná odpověď respondenta i na citlivou otázku je „ano“.

$p_{v,ano}$ = pravděpodobnost, že uslyšíme vynucenou odpověď „ano“.

$p_{v,ne}$ = pravděpodobnost, že uslyšíme vynucenou odpověď „ne“.

p_{ANO} = pravděpodobnost, že od respondenta i uslyšíme „ano“.

$$p_{ANO} = p_{v,ano} + (1 - p_{v,ano} - p_{v,ne}) \cdot a$$

Vynucená odpověď – vyhodnocení

Chceme zjistit (odhadnout) hodnotu a = pravděpodobnost, že správná odpověď respondenta i na citlivou otázku je „ano“.

$p_{V,ano}$ = pravděpodobnost, že uslyšíme vynucenou odpověď „ano“.

$p_{V,ne}$ = pravděpodobnost, že uslyšíme vynucenou odpověď „ne“.

p_{ANO} = pravděpodobnost, že od respondenta i uslyšíme „ano“.

$$p_{ANO} = p_{V,ano} + (1 - p_{V,ano} - p_{V,ne}) \cdot a$$

$$a = \frac{p_{ANO} - p_{V,ano}}{1 - p_{V,ano} - p_{V,ne}}$$

Vynucená odpověď – pizza

Hod'te kostkou. Pokud padne číslo 1, odpovězte „souhlasím“. Pokud padne číslo 6, odpovězte „nesouhlasím“. Pokud padne 2, 3, 4 nebo 5, vyjádřete svůj souhlas či nesouhlas s následujícím výrokem.

Ananas na pizzu patří.

Vynucená odpověď – pizza

Hod'te kostkou. Pokud padne číslo 1, odpovězte „souhlasím“. Pokud padne číslo 6, odpovězte „nesouhlasím“. Pokud padne 2, 3, 4 nebo 5, vyjádřete svůj souhlas či nesouhlas s následujícím výrokem.

Ananas na pizzu patří.

$$\hat{a} = \frac{3}{2} \left(\hat{p}_{ANO} - \frac{1}{6} \right)$$

Vynucená odpověď – pizza

Hod'te kostkou. Pokud padne číslo 1, odpovězte „souhlasím“. Pokud padne číslo 6, odpovězte „nesouhlasím“. Pokud padne 2, 3, 4 nebo 5, vyjádřete svůj souhlas či nesouhlas s následujícím výrokem.

Ananas na pizzu patří.

$$\hat{a} = \frac{3}{2} \left(\hat{p}_{ANO} - \frac{1}{6} \right)$$

Vyšel stejný vzorec, ale nemusel, záleží na zvolených pravděpodobnostech.

Pořád to není dost

Studie (1982), kde náhodný způsob volby mezi vynucenou a pravdivou odpovědí byl pod kontrolou experimentátorů: cca 25 % respondentů, kteří měli dát vynucenou odpověď „ano“, tak neučinilo.

Studie (1982), kde náhodný způsob volby mezi vynucenou a pravdivou odpovědí byl pod kontrolou experimentátorů: cca 25 % respondentů, kteří měli dát vynucenou odpověď „ano“, tak neučinilo.

Stydí se, i když tazatel neví, zda jde o pravdivou nebo vynucenou odpověď.

Pořád to není dost

Studie (1982), kde náhodný způsob volby mezi vynucenou a pravdivou odpovědí byl pod kontrolou experimentátorů: cca 25 % respondentů, kteří měli dát vynucenou odpověď „ano“, tak neučinilo.

Stydí se, i když tazatel neví, zda jde o pravdivou nebo vynucenou odpověď.

Slova „ano“ a „ne“ nahradíme neutrálními slovy, třeba „červená“ a „černá“.

Zneužívání systému sociálního zabezpečení (Nizozemsko, 2000):

I have two stacks of cards and a box behind in which I place the cards.
[GIVE THE BOX TO THE RESPONDENT AND LOOK AT IT
TOGETHER.] In the box, you find a card on which it is written what the
stack means: the right-hand stack is the 'yes' stack, and the left-hand
stack is the 'no' stack. [LET INTERVIEWEE LOOK AND GIVE
DIRECTIONS WITH THE NEXT EXPLANATION.]

Zneužívání systému sociálního zabezpečení (Nizozemsko, 2000):

I have two stacks of cards and a box behind in which I place the cards. [GIVE THE BOX TO THE RESPONDENT AND LOOK AT IT TOGETHER.] In the box, you find a card on which it is written what the stack means: the right-hand stack is the 'yes' stack, and the left-hand stack is the 'no' stack. [LET INTERVIEWEE LOOK AND GIVE DIRECTIONS WITH THE NEXT EXPLANATION.]

In the 'yes' stack [POINT TO THE RIGHT-HAND STACK] there are more red cards than in the 'no' stack [POINT TO THE LEFT-HAND STACK, RESPONDENT MAY CHECK]. If you want, you may shuffle the two stacks [SEPARATELY].

Maskovaná odpověď (1990)

Zneužívání systému sociálního zabezpečení (Nizozemsko, 2000):

I have two stacks of cards and a box behind in which I place the cards. [GIVE THE BOX TO THE RESPONDENT AND LOOK AT IT TOGETHER.] In the box, you find a card on which it is written what the stack means: the right-hand stack is the 'yes' stack, and the left-hand stack is the 'no' stack. [LET INTERVIEWEE LOOK AND GIVE DIRECTIONS WITH THE NEXT EXPLANATION.]

In the 'yes' stack [POINT TO THE RIGHT-HAND STACK] there are more red cards than in the 'no' stack [POINT TO THE LEFT-HAND STACK, RESPONDENT MAY CHECK]. If you want, you may shuffle the two stacks [SEPARATELY].

Now, please take from each stack an arbitrary card. You may take the card on top or from within the stack. [TAKE A CARD FROM EACH STACK] Nobody but you can see the colors of your cards; when you mention a card color, we do not know the stack from which you took the card. Thus, your privacy is guaranteed: your answer will always remain a secret. [...] I propose that we now try out a few questions.

Then, respondents answer with “red” or “black” to the set of questions, which include the following:

Then, respondents answer with “red” or “black” to the set of questions, which include the following:

At a social services check-up, have you ever acted as if you were sicker or less able to work than you actually are? Have you ever noticed an improvement in the symptoms causing your disability, for example in your present job, in volunteer work, or the chores you do at home, without informing the Department of Social Services of this change?

Maskovaná odpověď – vyhodnocení

Chceme zjistit (odhadnout) hodnotu a = pravděpodobnost, že správná odpověď respondenta i na citlivou otázku je „ano“, tedy že říká barvu vylosovanou z balíčku „ano“.

Maskovaná odpověď – vyhodnocení

Chceme zjistit (odhadnout) hodnotu a = pravděpodobnost, že správná odpověď respondenta i na citlivou otázku je „ano“, tedy že říká barvu vylosovanou z balíčku „ano“.

p_A = pravděpodobnost, že z balíčku „ano“ vylosuji červenou kartu.

Maskovaná odpověď – vyhodnocení

Chceme zjistit (odhadnout) hodnotu a = pravděpodobnost, že správná odpověď respondenta i na citlivou otázku je „ano“, tedy že říká barvu vylosovanou z balíčku „ano“.

p_A = pravděpodobnost, že z balíčku „ano“ vylosuji červenou kartu.

$1 - p_A$ = pravděpodobnost, že z balíčku „ne“ vylosuji červenou kartu.

Maskovaná odpověď – vyhodnocení

Chceme zjistit (odhadnout) hodnotu a = pravděpodobnost, že správná odpověď respondenta i na citlivou otázku je „ano“, tedy že říká barvu vylosovanou z balíčku „ano“.

p_A = pravděpodobnost, že z balíčku „ano“ vylosuji červenou kartu.

$1 - p_A$ = pravděpodobnost, že z balíčku „ne“ vylosuji červenou kartu.

p_{RED} = pravděpodobnost, že od respondenta i uslyšíme „červená“.

Maskovaná odpověď – vyhodnocení

Chceme zjistit (odhadnout) hodnotu a = pravděpodobnost, že správná odpověď respondenta i na citlivou otázku je „ano“, tedy že říká barvu vylosovanou z balíčku „ano“.

p_A = pravděpodobnost, že z balíčku „ano“ vylosuji červenou kartu.

$1 - p_A$ = pravděpodobnost, že z balíčku „ne“ vylosuji červenou kartu.

p_{RED} = pravděpodobnost, že od respondenta i uslyšíme „červená“.

$$p_{RED} = a \cdot p_A + (1 - a) \cdot (1 - p_A)$$

Maskovaná odpověď – vyhodnocení

Chceme zjistit (odhadnout) hodnotu a = pravděpodobnost, že správná odpověď respondenta i na citlivou otázku je „ano“, tedy že říká barvu vylosovanou z balíčku „ano“.

p_A = pravděpodobnost, že z balíčku „ano“ vylosuji červenou kartu.

$1 - p_A$ = pravděpodobnost, že z balíčku „ne“ vylosuji červenou kartu.

p_{RED} = pravděpodobnost, že od respondenta i uslyšíme „červená“.

$$p_{RED} = a \cdot p_A + (1 - a) \cdot (1 - p_A)$$

$$a = \frac{1}{2p_A - 1} (p_{RED} + p_A - 1)$$

Nic není zadarmo, kde je tedy cena, kterou platíme?

Nic není zadarmo, kde je tedy cena, kterou platíme?

Vyšší variabilita odhadu, vnášíme do problému dodatečný šum.

Nic není zadarmo, kde je tady cena, kterou platíme?

Vyšší variabilita odhadu, vnášíme do problému dodatečný šum.

Jsou dostupné i další techniky pro zjišťování binárních odpovědí.

Nic není zadarmo, kde je tedy cena, kterou platíme?

Vyšší variabilita odhadu, vnášíme do problému dodatečný šum.

Jsou dostupné i další techniky pro zjišťování binárních odpovědí.

Jsou dostupné i techniky pro zjišťování kvantitativních odpovědí.