

Populacje „ukryte” – rekonstrukcja
populacji na przykładzie metody
Respondent Driven Sampling

Populacje „ukryte” – dobór próby sterowany przez respondentów

Populacja „ukryta” to populacja, której jednostki z różnych powodów nie są uwzględniane w oficjalnych statystykach, nie istnieje operat losowania dla takiej populacji

Przykłady badań populacji ukrytych badanych za pomocą RDS

Osoby chore/zarażone AIDS/HIV – badania prowadzone w trzech miastach stanu Connecticut, University of Connecticut

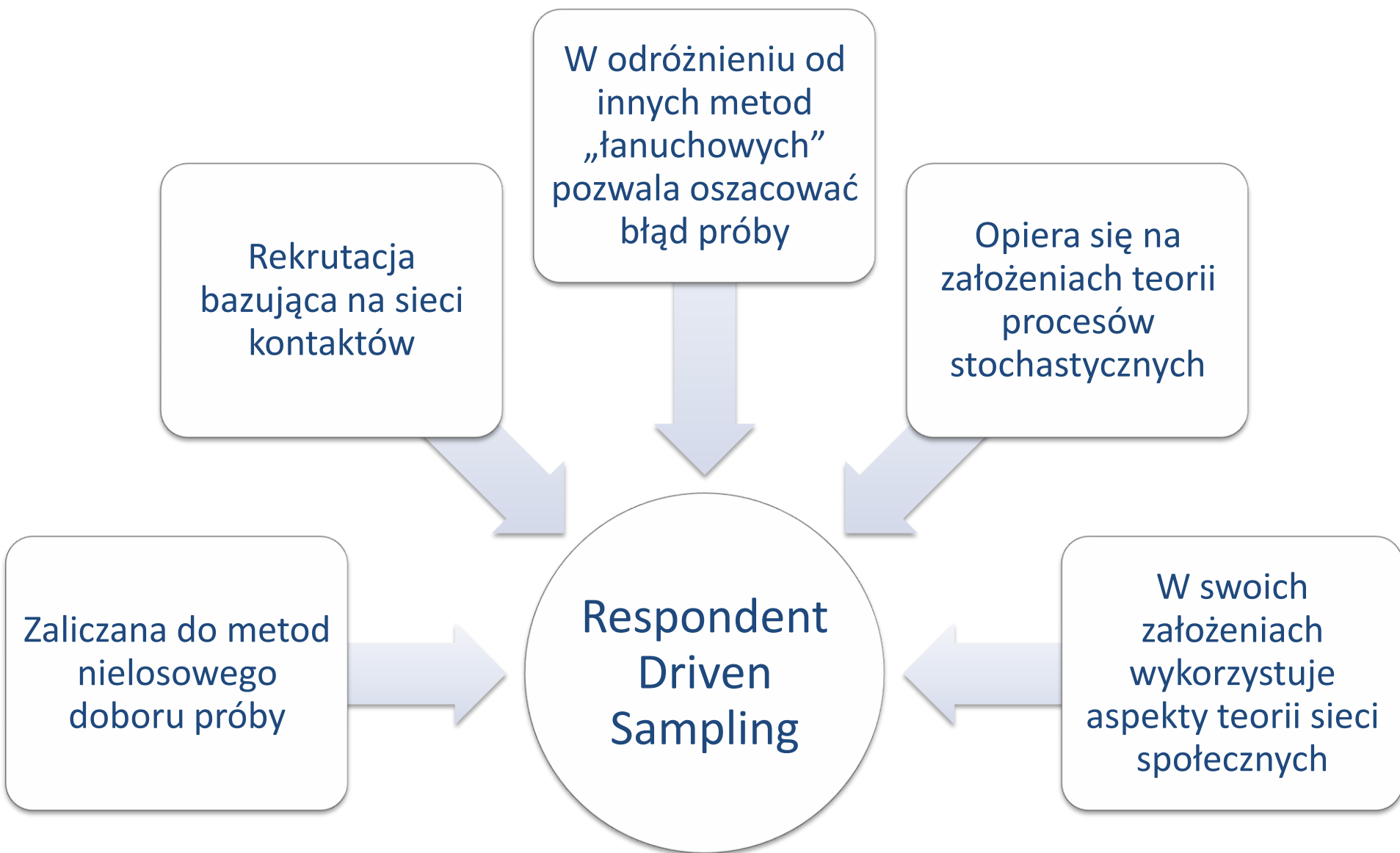
Uzależnieni od narkotyków – badania prowadzone w Tijuanie, American Journal of Addiction

Muzycy jazzowi – badanie przeprowadzone w Nowym Jorku, Cornell University

Imigranci – Badania prowadzone w Warszawie, Ośrodek Badań nad Migracjami

Osoby homoseksualne i transseksualne – badanie przeprowadzone w Kapsztadzie, Medical Research Council

Respondent Driven Sampling



Respondent Driven Sampling - założenia

1. Osoba będąca częścią populacji ukrytej zna innych członków tej populacji i są oni dla niej dostępni(kontakt)

2. Każdy respondent jest osiągalny przez innego respondenta poprzez sieć kontaktów

3. Respondenci pojawiają się tylko raz w próbie

4. Respondenci są w stanie określić swoją sieć kontaktów(wielkość sieci i powiązań)

5. Respondenci rekrutują swoich znajomych w sposób przypadkowy.

Respondent Driven Sampling – jak przebiega badanie

1. Wybór „ziarenek”

2. Zrekrutowanym „ziarenkom” proponujemy nagrodę, które pozwolą na dalszą rekrutację poprzez wręczone kupony

3. Zrekrutowani przez „ziarenka” również otrzymują nagrodę za udział w badaniu i dalsza rekrutację

4. Stosowanie systemu sprawdzania przynależności do populacji ukrytej

5. Identyfikacja sieci znajomości zrekrutowanych

6. Dodatkowe nagrody na zrekrutowanie trudniej dostępnych osób

7. Badanie może zostać zakończone w dwóch przypadkach, przy dostatecznie dużej saturacji populacji, lub przy osiągnięciu zbieżności z zakładaną populacją.

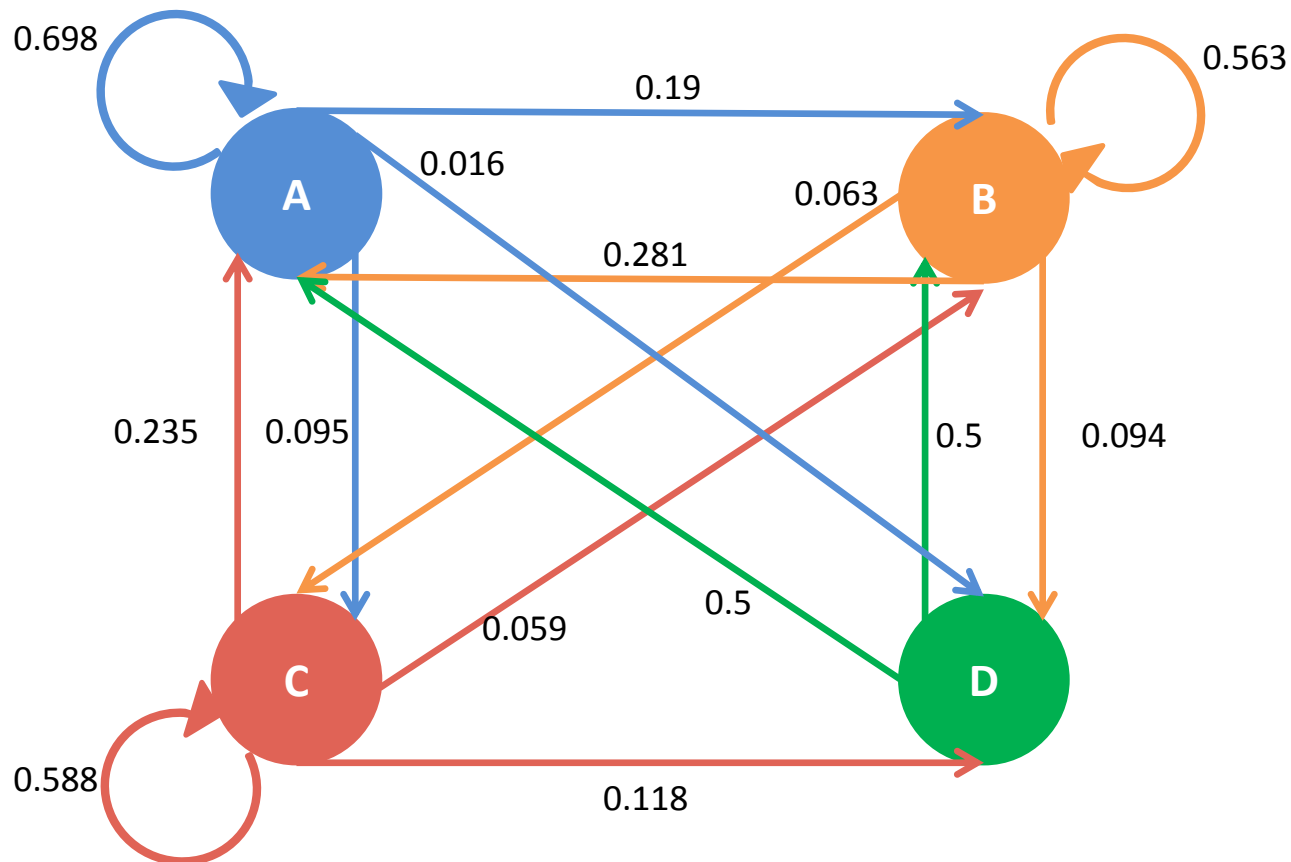
Respondent Driven Sampling – proces stochastyczny

Po przeprowadzeniu badania wstępnego według wymienionych wcześniej założeń uzyskano następujący rozkład:

Rekrutujący	Zrekrutowany				
	A	B	C	D	Total
A	69.8%	19%	9.5%	1.6%	100% (63)
B	28.1%	56.3%	6.3%	9.4%	100% (32)
C	23.5%	5.9%	58.8%	11.8%	100% (17)
D	50%	50%	0%	0%	100% (4)
Rozkład zrekrutowanych	50.9% (59)	28.4% (33)	15.5% (18)	5.2% (6)	100% (116)

Respondent Driven Sampling – proces stochastyczny

Zatem respondenci rekrutowali siebie wzajemnie w według schematu:



Respondent Driven Sampling – proces stochastyczny

Równowaga wynikająca z układu równań:

$$1 = E_A + E_B + E_C + E_D$$

$$E_A = 0.698 * E_A + 0.281 * E_B + 0.235 * E_C + 0.5 * E_D$$

$$E_B = 0.19 * E_A + 0.563 * E_B + 0.059 * E_C + 0.5 * E_D$$

$$E_C = 0.095 * E_A + 0.063 * E_B + 0.588 * E_C$$

$$E_A = 0.490$$

$$E_B = 0.297$$

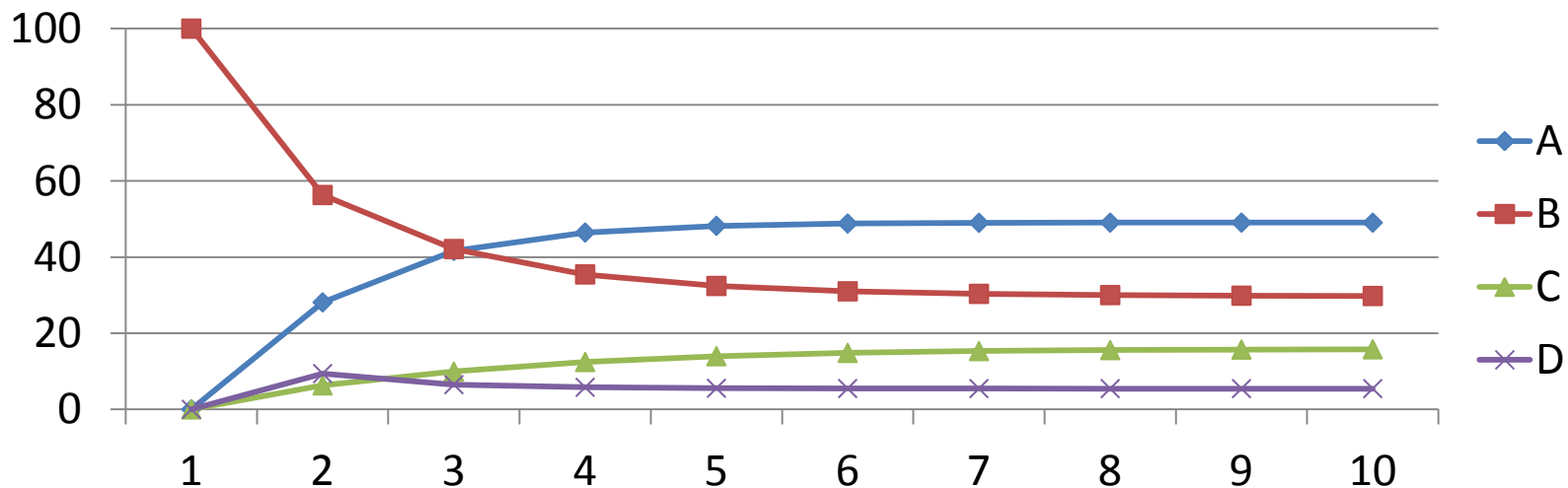
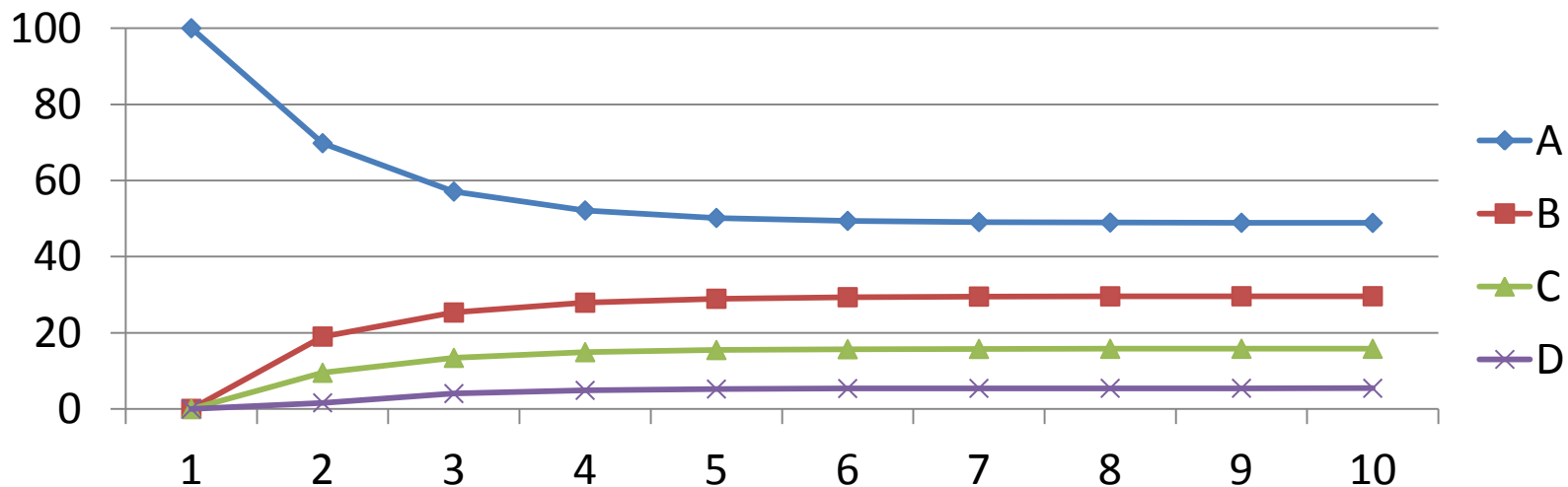
$$E_C = 0.158$$

$$E_D = 0.054$$

	A	B	C	D	Total
Rozkład zrekrutowanych	50.9% (59)	28.4% (33)	15.5% (18)	5.2% (6)	100% (116)
Równowaga	49%	29.7%	15.8%	5.4%	100%

Respondent Driven Sampling – proces stochastyczny

Równowaga uzyskiwana w kolejnych krokach przy różnym wyborze „ziarenek”:



Respondent Driven Sampling – problemy

Co zrobić z respondentami przed osiągnięciem przez próbę równowagi.

Skomplikowane wyliczenia błędów

Błędy płynące z dużych indywidualnych sieci kontaktów,

Jak zmotywować respondentów do rekrutacji