## Three sides of the same coin: Teaching, Research and Learning





Personalized Treatment: Sounds heavenly, but where on Earth can they find the right guinea pig for me?

> Keli Liu and Xiao-Li Meng & the "Happy Team" Department of Statistics Harvard University

## A trio of papers

- Meng, X.-L. (2014). A Trio of Inference Problems that Could Win You a Nobel Prize in Statistics (If You Help Fund It). *In Past, Present, and Future of Statistical Science (Eds Lin et. al.)*. ( <u>http://www.stat.harvard.edu/Faculty\_Content/meng/COPSS\_50.pdf</u>)
- Liu, K and Meng, X.-L. (2014). A Fruitful Resolution to Simpson's Paradox via Multi-Resolution Inference. *The American Statistician, Vol* 68, pp 17-29. (<u>http://statistics.fas.harvard.edu/people/xiao-li-meng</u>)
- Liu, K. and Meng, X.-L. (2015). There is individualized treatment. Why not individualized inference? To appear in *Annual Review of Statistics and Its Applications*. (available by emailing meng@stat.harvard.edu)

#### EM 16: Real-Life Statistics: Your Chance for Happiness (or Misery)



#### Graduate Seminar for General Education (GSGE)

- Graduate Courses designed for designing Gen Ed courses
- Taken by both graduate students and undergraduate students
- An ideal way of integrating research and pedagogy
- \*A great venue for training team work skills
- A unique learning and re-education experience for faculty

Graduate Seminar for Undergraduate Education (GSUE)
Strend GSGE for designing any undergraduate courses

## What does Big Data mean for you? We see you and others more clearly



### Gift: Treatment for you based only on data from people like you. Curse: No one is perfectly like YOU.



## **A Painful Problem**



Ms.Payne



#### **Kidney Stone Treatment** C. R. Charig, D. R. Webb, S. R. Payne, O. E. Wickham (March 1986)

Br Med J (Clin Res Ed) 292 (6524): 879-882.

Treatment A	Treatment B
78%	83%
(273/350)	(289/350)

	Treatment A	Treatment B
Small	<b>93%</b>	87%
Stone	(81/87)	(234/270)
Large	<b>73%</b>	69%
Stone	(192/263)	(55/80)

**A: Open Surgery; B: Percutaneous Nephrolithotomy** 7/25/16

## Simpson's Paradox

## **Trivial Mathematics**

It is possible for

$$\frac{A_1}{C_1} > \frac{A_2}{C_2}$$
 and  $\frac{B_1}{D_1} > \frac{B_2}{D_2}$ 

BUT

$$\frac{A_1 + B_1}{C_1 + D_1} < \frac{A_2 + B_2}{C_2 + D_2}$$

















Uneven distribution of stone sizes across treatments makes overall success rate misleading.

# And it is not just for medical treatments ....

Another large-profile case of Simpson's paradox was described in "*Sex Bias in Graduate Admissions: Data from Berkeley.*" (P.J. Bickel, E.A. Hammel and J.W. O'Connell). *Science* **187** (4175): 398–404).

In this case, the University of California was worried about being sued for bias against women applying to graduate school. The admission figures for Fall 1973 showed that men applying were more likely than women to be admitted, and the difference was so large that it was unlikely to be due to chance.

## Berkeley Sex Bias Case ...

	Applicants	Admitted
Men	8442	44%
Women	4321	35%

## **Case Closed !!!**

## **NOT SO FAST!**

Major	Men		Women	
	Applicants	Admitted	Applicants	Admitted
А	825	62%	108	<u>82%</u>
В	560	63%	25	<u>68%</u>
С	325	<u>37%</u>	593	34%
D	417	33%	375	<u>35%</u>
E	191	<u>28%</u>	393	24%
F	272	6%	341	<u>7%</u>
Total	2590	46%	1835	30%

Out of 101 departments, only 10 were significant at 0.05 level, with 6 higher on female and 4 higher on male.

# Who is Simpson?



#### Homer: "Oh, people can come up with statistics to prove anything. 14% of people know that."





#### **The Real Simpson**

Edward Hugh Simpson is a British statistician who described this phenomenon in his paper: "The Interpretation of Interaction in Contingency Tables." *Journal of the Royal Statistical Society, Ser. B* **13**: 238–241. (1951)

#### Edward Hugh Simpson (1922 - )



#### A different scenario but with the same numbers Clinical Trial for Treating Pain

Treatment A	Treatment B
78%	<b>83%</b>
(273/350)	(289/350)

	Treatment A	Treatment B
Small	<b>93%</b>	87%
Stone	(81/87)	(234/270)
Large	<b>73%</b>	69%
Stone	(192/263)	(55/80)

Sizes were measured after applying treatments for pain

• Treatment A has no effect on the size

## Remember...



Ms.Payne



## Ignore personal information?

Treatment A	Treatment B
78% 🔶	<mark>→</mark> 83%
(273/350)	(289/350)



**Resolution 0** 

	TreatmentA	Treatment B
Small	<b>93%</b>	87%
Stone	(81/87)	(234/270)
Large	<b>73%</b>	69%
Stone	(192/263)	(55/80)

## Incorporate personal information?

Treatment A	Treatment B
78% 🔶	→ 83%
(273/350)	(289/350)



**Resolution 0** 

	TreatmentA	Treatment B
Small	<b>93%</b>	87%
Stone	(81/87)	(234/270)
Large	<b>73% ←</b>	→ 69%
Stone	(192/263)	(55/80)



**Resolution 1** 

• But did we use the information correctly?

### **Deceptive But False Similarity**

> Are Mark and Ben similar?



What if side effect of treatment is hair loss and the above are post-treatment pictures?





## So how about something in between?

Trea	atmentA	Trea	atment B		
7	/8% 🔶	> 83%			$\longleftrightarrow$
(27	3/350)	(28	9/350)		Resolution 0
	Treatment A	,	Treatment B	;	$\leftarrow \rightarrow$
		<u>.</u>			Resolution 0.5
	93%		87%		
;	(81/87)		(234/270	)	$\longleftrightarrow$
	73% <		→ 69%		Resolution 1
)	(192/263	)	(55/80)		
	Trea 7 (27	Treatment A 78% <	Treatment A       Treatment A         78% ←       8         (273/350)       (28)         Treatment A       (28)         93%       (28)         (81/87)       (28)         73% ←       (192/263)	Treatment A       Treatment B $78\%$ $83\%$ $(273/350)$ $(289/350)$ (273/350) $(289/350)$ Treatment A       Treatment B         93% $87\%$ $(81/87)$ $87\%$ $(192/263)$ $69\%$ $(192/263)$ $(55/80)$	Treatment A       Treatment B $78\% \leftarrow 83\%$ $(273/350)$ $(289/350)$ Treatment A       Treatment B $93\%$ 87% $(81/87)$ 87% $(192/263)$ $(55/80)$

## A Multi-Resolution View of Big Data

#### Population Resolution



#### Individual Resolution



## The Original Happy Team



7/25/16

## **The Extended Happy Team**



## **Movie Time!**