Comments on Science, Experimentation and C2 Research

10th ICCRTS

Ritz-Carlton

McLean, VA

Dr. Paul J. Hiniker
Sr. Ops Research
June 14, 2005

How do we determine causality?

- "Development of Western science is based on two great achievements:
  - The invention of the formal logical system (in Euclidean geometry) by the Greek philosophers, and
  - the discovery of the possibility to find out causal relationships by systematic experiment (during the Renaissance)."

(Albert Einstein, 1953)

- The causal interpretation of a simple(or partial) correlation depends upon
  - the presence of a compatible causal hypothesis (H_c)
  - and the absence of a plausible rival hypothesis (H_r) to explain the correlation on other grounds.

(Herb Simon, 1957)

Decomposing Simple Correlations with Controls, e.g. Incendiary Fire Engines

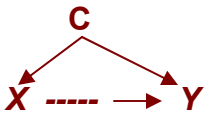
X --→ Y

Hc:Fire engines prevent fire damage.

	not X	X	
Y	.20	.30	.50
notY	.30	.20	.50
	.50	.50	100 OBSERVATIONS

X =Many Fire Engines Sent

Y =Much Fire Damage at Site



Hr:Big fires cause more fire damage.

(1) Simple Correlation, ***n = 100 Observations***

$$\Phi_{XY} = P_{XY} - P_X P_Y \,/\, \sqrt{(P_X Q_X P_Y Q_Y)} = .20!$$

(2) Yule's Covariance Theorem

C=Size of Fire>1000ft²

Partial correlations(controlling for c), ***n=50***

$\Phi_{XY\text{-}C} = -.25$

$\Phi_{XYC} = -.25$

$\Phi_{XC} = .60$

$\Phi_{YC} = .60$

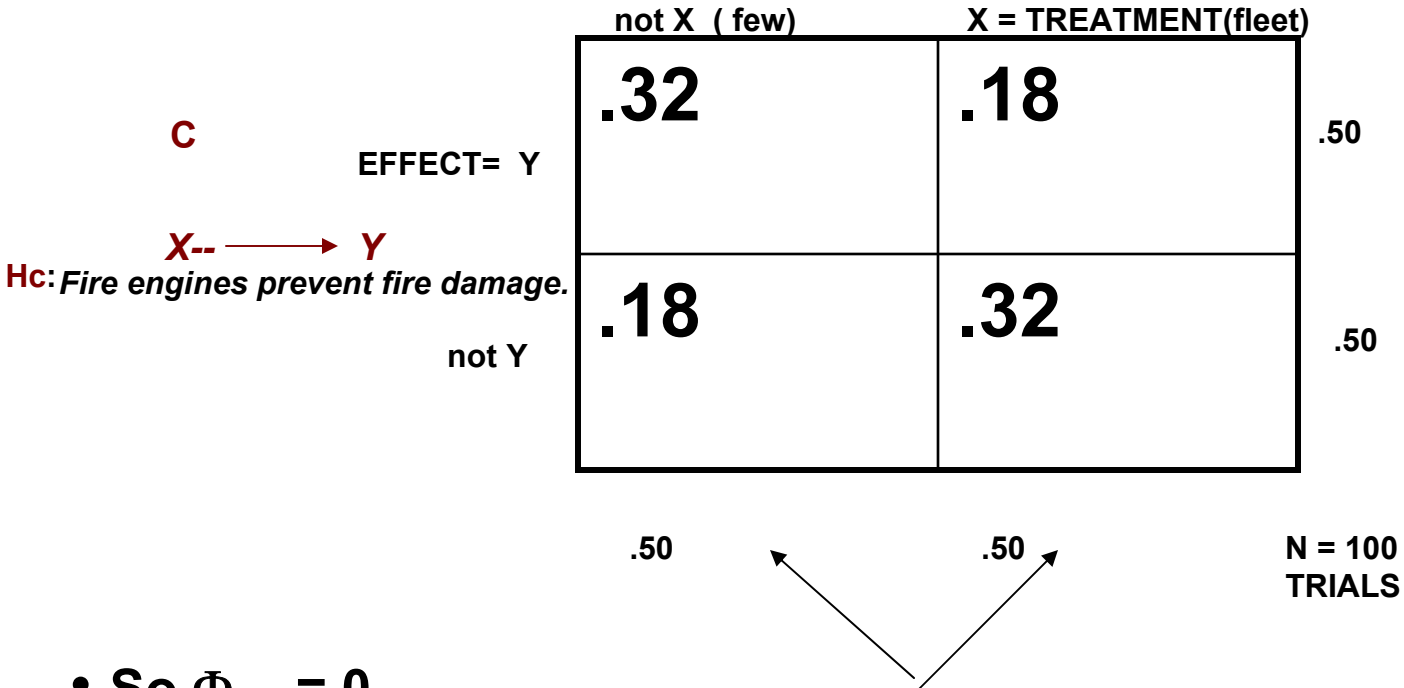
$$\Phi_{XY} =$$

$$\Phi_{XY\text{-}C} P_{\text{-}C} \sqrt{(P_{X\text{-}C}\, Q_{X\text{-}C} P_{Y\text{-}C} Q_{Y\text{-}C} / P_X Q_X P_Y Q_Y)}$$

$$+ \Phi_{XYC} P_C \sqrt{(P_{XC} Q_{XC} P_{YC} Q_{YC} / P_X Q_X P_Y Q_Y)}$$

$$+ \Phi_{YC} \Phi_{XC}$$

Controlled Experiment, e.g. Fire Engines Prevent Fire Damage



- So $\Phi_{XC} = 0$.
- Since $\Phi_{XC} = 0$,
  Experiment $\Phi_{XY} = (w' \Phi_{XY\text{-}C} + w \Phi_{XYC})/ 2 = \Phi_{XY.C}$, for $\forall c$,
  thus ruling out rival explanations, e.g. size of fire
- Experiment $\Phi_{XY} = -.28$
- So less fire damage is due to more fire engines on site.

Causal Modeling with Non-Experimental Data

- So to prevent spurious correlation, conduct of a controlled experiment guarantees $ø_{cx} = 0$ and ensures a valid test of a causal hypothesis.

- However, for non-experimental causal modeling,with one or more independent variables, one must verify that the residual error terms of all the variables are uncorrelated:

$$r_{Uy\,Uxi} = r_{Uxi\,Uxj} = 0, \text{ for all } X_i.$$

  - Otherwise, there could exist some extraneous variable(s), C_i, affecting both Y and X_i, hence forming part of u_y and u_{xi}, which would then be correlated; this would spuriously contribute to the correlations implied by the model. (Simon, 1957)

- Thus simple correlation can neither prove nor disprove a causal hypothesis; and controlled experimentation can. (Shadish, Cook, & Campbell, 2002)

  - Uncontrolled discovery trials can suggest useful new hypotheses; but they cannot prove or disprove a causal hypothesis. (Pearl, 2000; GUIDEx, 2005)