Causal Inference and Invariance

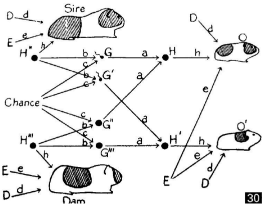
Charles Zheng and Qingyuan Zhao

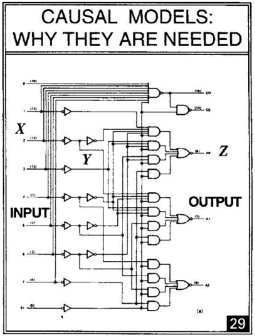
Stanford University

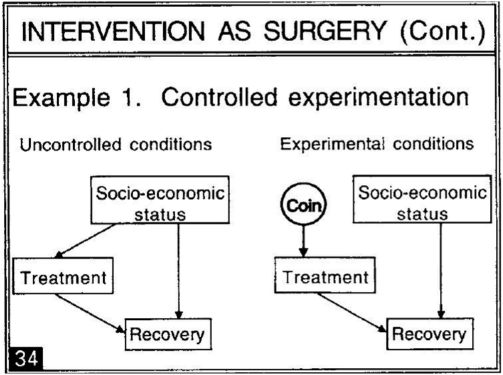
February 13, 2016

(Part 1/2)

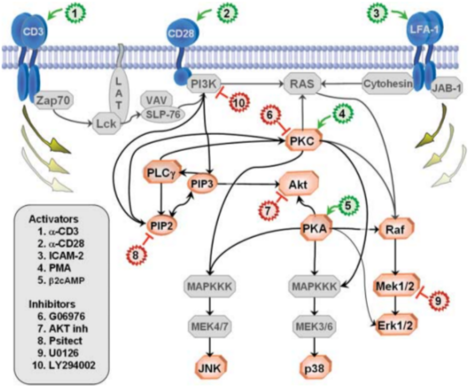
Understanding = cause and effect





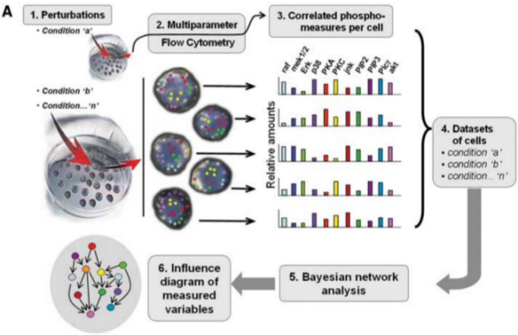


A hot application: systems biology



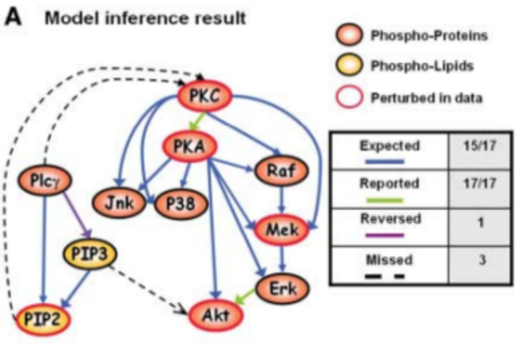
- Causal relationships = *chemical interactions*.
- Experimenters *intervene* by injecting *activators* and *inhibitors*.

Protein signalling data



- Flow cytometry data from Sachs et al.*Science*, 2005.
- 1 observational data set + 9 interventions.

Putative causal model



- Causal inference applied to observational + interventional data.
- Recovered most of the known interactions.

The many facets of causality

- *Philosophy.* What is causality? How do we learn about cause and effect? *Aristotle, Hume.*
- *Computer science.* Can we build an artificial intelligence which reasons like humans? *Judea Peal.*
- *Social science.* What influences an individual's life choices?
- *Law.* Whose "fault" is it??
- *Statistics.* Answering the above questions using data!

Statistics and causality

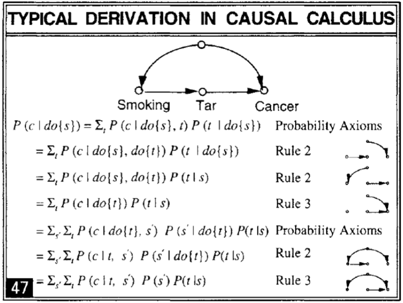
- *Estimating causal effects from data.* Can we predict a causal effect based on observational or experimental data? E.g. effect of a medical treatment based on clinical trial data? Motviation for potential outcomes approach developed by Rubin, etc.
- *Bayesian networks/structure learning from data.* Can we model multivariate relationships using a network structure? Networks *can be* given causal interpretation, but causal inference is not the only motivation. Motivation for graphical lasso.

Section 2

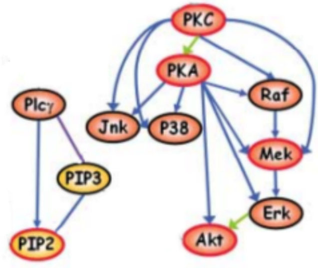
Principles of Causal Inference

Principles of Causal Inference

Best framework: *Graphical approach* pioneered by Judea Pearl.

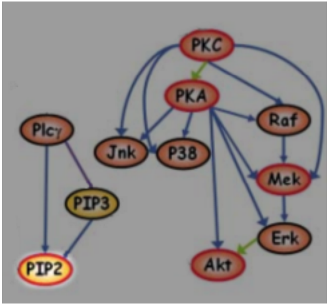


Graphs: nodes and vertices



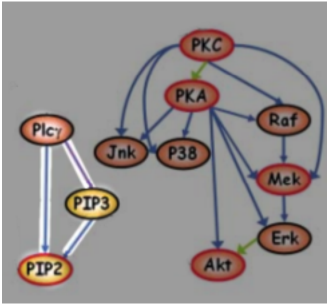
- Each variable in the dataset is given a *node*.
- Arrows indicate which variables *cause* which other variables. (Parents $\rightarrow$ children).
- Undirected or bidirected edges = correlation due to mutual causation or latent common causes.

Causality and experiments



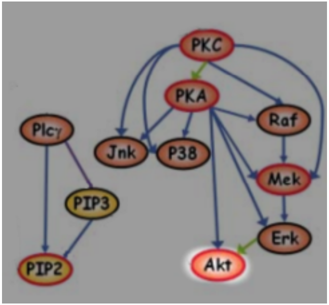
Intervening on variables in the system causes the distribution to change.

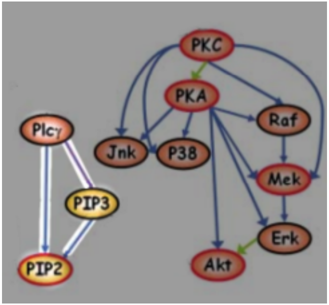
Causality and experiments



- Not every variable will be affected by the intervention!
- Following the arrows tells you *which* variables which are affected.

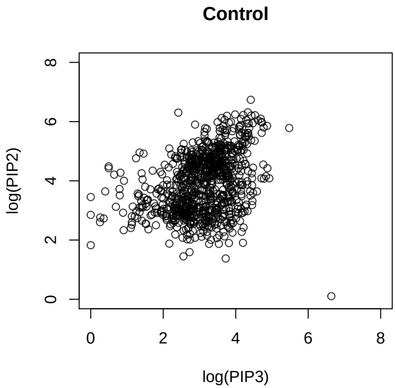
Principle I: Which variables are affected.





- If we *inhibit* Akt, no other variables should be affected.
- If we *inhibit* PIP2, then we may not only change the distribution of PIP2, but also PIP3.

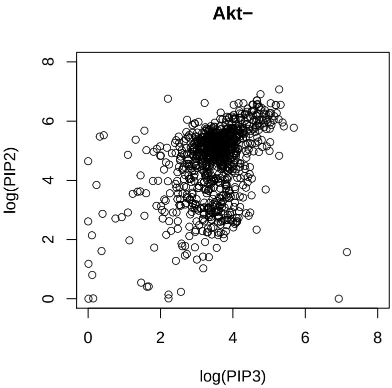
Principle I: Which variables are affected.



Looking at Sachs data.

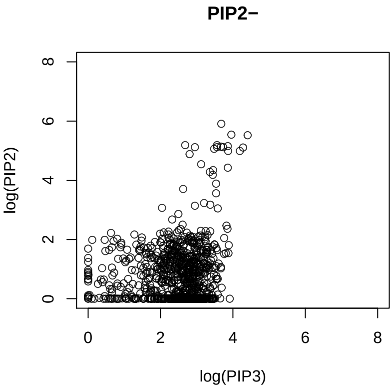
Joint distribution of PIP2 and PIP3 in the “control” case.

Principle I: Which variables are affected.



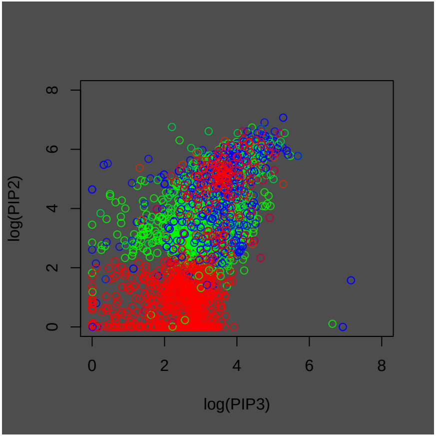
Joint distribution of PIP2 and PIP3 when we intervene on Akt.

Principle I: Which variables are affected.



Joint distribution of PIP2 and PIP3 when we intervene on PIP2.

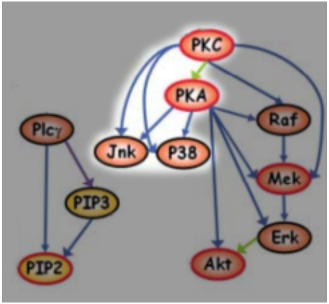
Principle I: Which variables are affected.



Control, **PIP2-**, **Akt-**

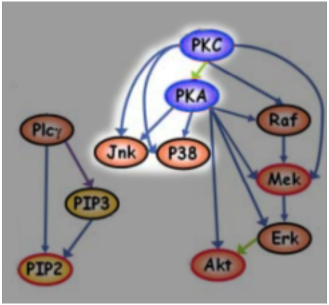
Intervening on PIP2 also affects the distribution of PIP3, while intervening on Akt does not (drastically) change the distribution.

Principle II: Conditional independence.



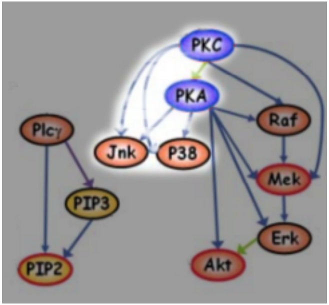
- Surprisingly, the structure of the causal graph implies certain *conditional independence* relationships.
- This allows the potential to infer causal relationships from observational data.

Principle II: Conditional independence.



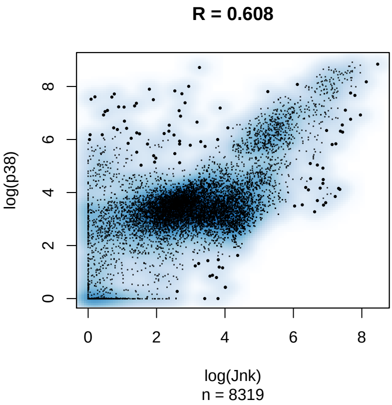
- Two variables are independent conditional on their common parents.
- Conditioning on PKC and PKA, Jnk and p38 should be independent.

Principle II: Conditional independence.



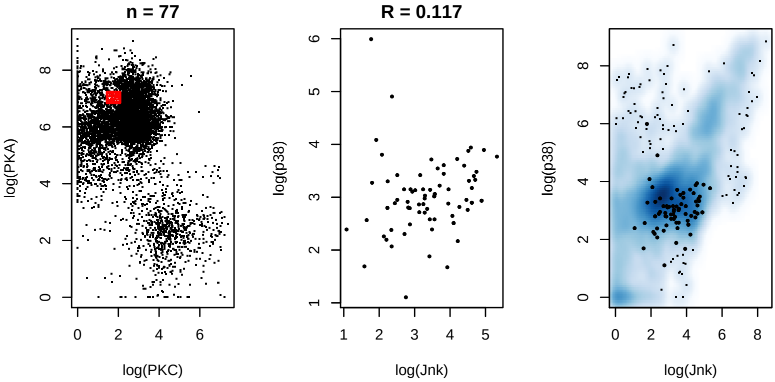
- "Once you and I condition on common factors, we are left with nothing in common."

Principle II: Conditional independence.



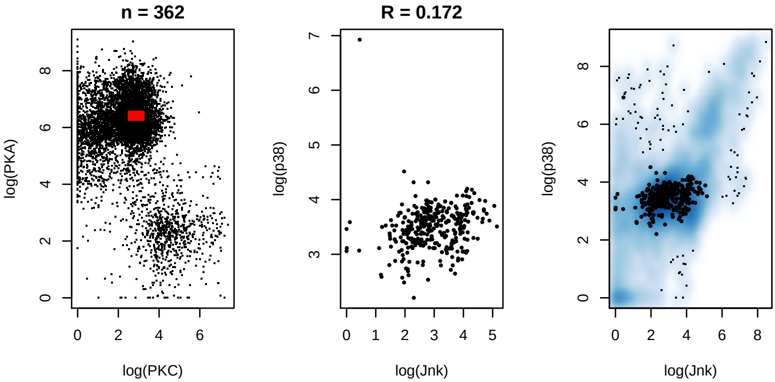
Marginally, p38 and Jnk are correlated.

Principle II: Conditional independence.



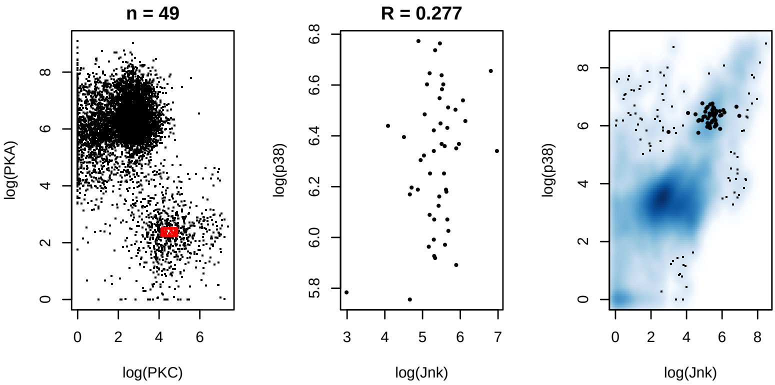
We can't condition on PKA and PKC since the data is continuous. But, conditioning on small windows seems to reduce association.

Principle II: Conditional independence.



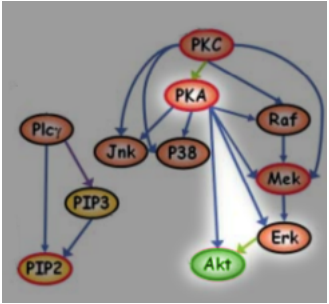
Left: We condition on (PKA, PKC) to lie within the indicated window. *Center:* Conditional joint distribution of (Jnk, p38). *Right:* Conditional join distribution, overlaid on marginal distribution.

Principle II: Conditional independence.



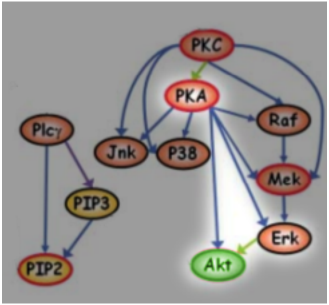
PKA and PKC *explain away* some (if not all) of the association between Jnk and p38. (Recall that R = 0.608 marginally.)

Principle III: Predictive invariance



- The conditional distribution $\Pr[Akt|PKA, Erk]$ is invariant to interventions applied to other variables.
- Therefore, the optimal rule for predicting $\hat{Akt}(PKA, Erk)$ is invariant as well.

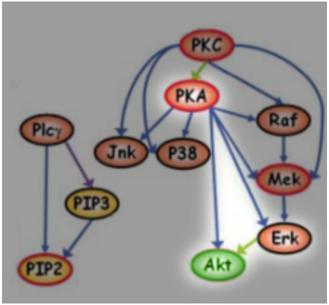
Principle III: Predictive invariance

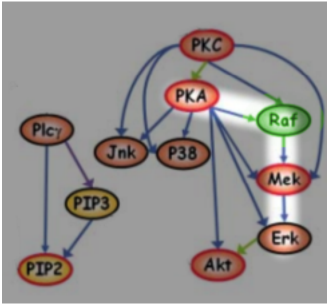


$\{PKA, Erk\}$ is an "invariant set" for Akt since:

- It includes all of the "direct" causes of Akt in the graph.
- It doesn't include any variables caused by Akt.

Principle III: Predictive invariance

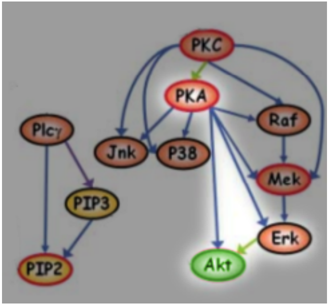


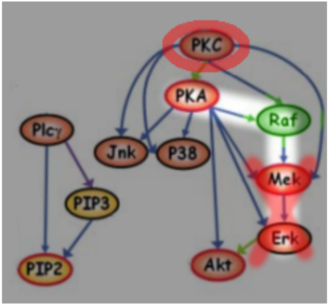


In contrast, $\{PKA, Mek, Erf\}$ is *not* an invariant set for *Raf* since:

- .
- .

Principle III: Predictive invariance

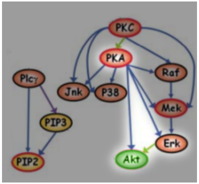


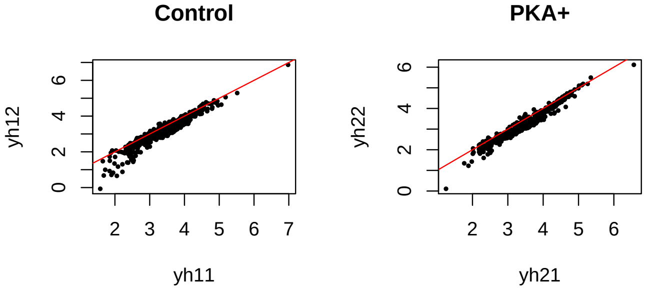


In contrast, $\{PKA, Mek, Erf\}$ is *not* an invariant set for *Raf* since:

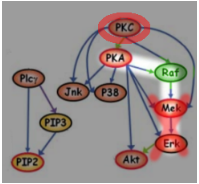
- It is missing a direct cause of *Raf*.
- It contains variables which are caused by *Raf*.

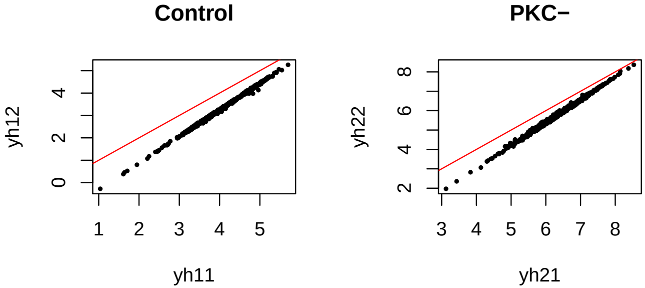
$\{PKA, Erk\}$ is an invariant set for Akt.





$\{PKA, Mek, Erf\}$ is *not* an invariant set for *Raf*.





Overview: Principles of Causal Inference

Causal relationships in a system represented by a graph. The graph tells you:

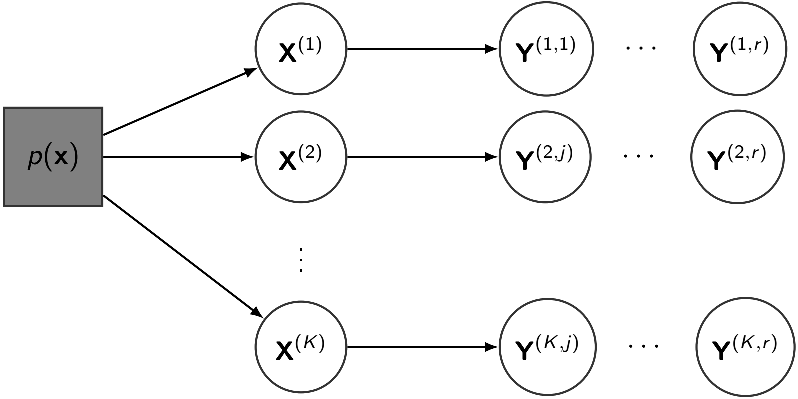
I. which variables are affected by an intervention.

II. what conditional independence relationships exist in the joint distribution.

III. which sets of predictors and responses will have "invariant" optimal predictive rules.

Look! A diagram!

Don't put this in the final presentation.



Legend: $K = \{$ **2**, **9**, **99**, **999** $\}$