

Bayesian graphical modelling in genetic epidemiology

- Causal inference: classical and Bayesian approaches
- Exploiting genotypes as instrumental variables: Mendelian randomization
- Limits to Mendelian randomization
- More general methods for reverse-engineering genotype-phenotype relationships
 - Sparse Bayesian instrumental variable analysis

Practical exercises

- Basic principles of Bayesian hypothesis testing
- Using VIBES (variational Bayes package) to evaluate the evidence for different models
- Using JAGS (Markov chain Monte Carlo package) for instrumental variable analysis with genetic instruments
- Using SPIV (sparse Bayesian instrumental variable package) with genome-wide data to infer causal relationships between biomarkers and outcome

Using genetic variation to infer causal biomarker-disease associations

Bayesian instrumental variable analysis

- “-omic” epidemiology yields many phenotypic *biomarkers* that predict outcome
 - metabolic measurements, gene expression levels, serum proteins
- We want to infer which biomarker-disease associations are causal
 - possible therapeutic targets
 - as surrogate end-points in early-stage clinical trials

Using genetic variation to infer causality in phenotypic biomarker-disease associations

- Genotypes are randomized at meiosis
 - if population stratification is controlled, associations with phenotype are unconfounded except by short-range allelic associations
- Can we exploit genotypic effects on phenotypic biomarkers to infer causal relationships between biomarkers and outcome?
 - infer mechanisms of genotypic effects
 - predict therapeutic or adverse effect of intervention on a pathway of interest
 - validate surrogate end-points for clinical trials

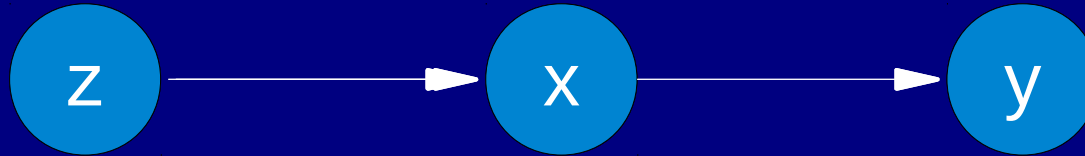
Classical approaches to causal inference

- Experimentalists: causal relationships can be inferred only from randomized intervention
- *Structural causal model* (Pearl): causation can be inferred if one of three conditions holds
 - an instrumental variable has been measured (randomization is a special case)
 - all confounders have been measured (back-door criterion)
 - an unconfounded variable on the causal pathway has been measured (front-door criterion)

Directed acyclic graphs

- nodes connected to each other by directed edges, with no cycles
- When the edges define probabilistic dependencies, a DAG is a Bayesian network
- *Markov blanket* of a node consists of the “parents” and “children” of that node
- Inference methods for Bayesian networks are not necessarily Bayesian
 - updating of nodes uses Bayes theorem

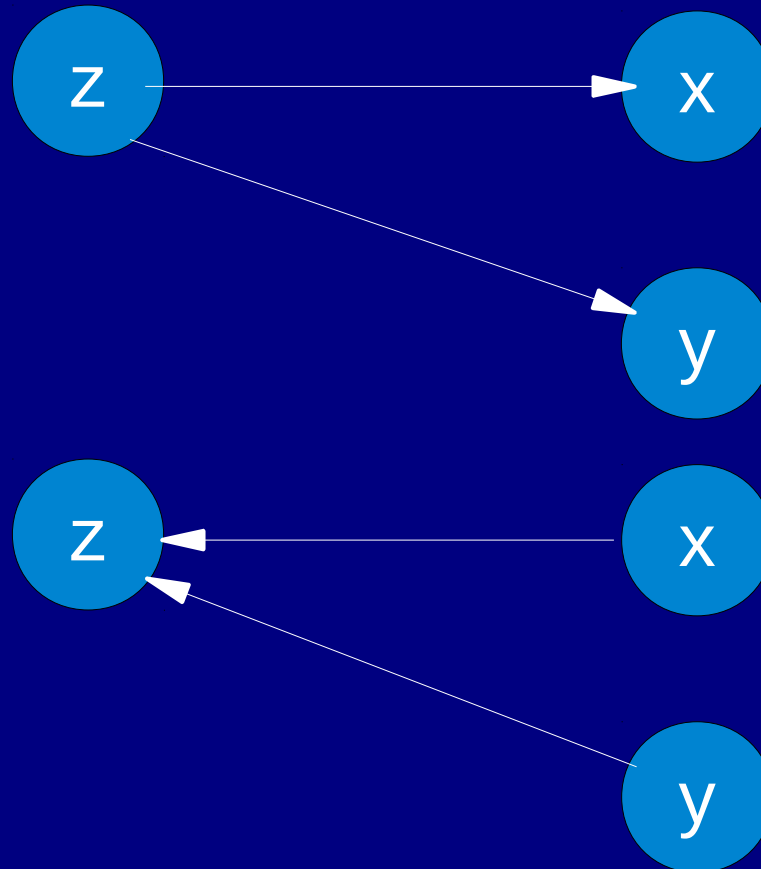
Conditional independence in graphs



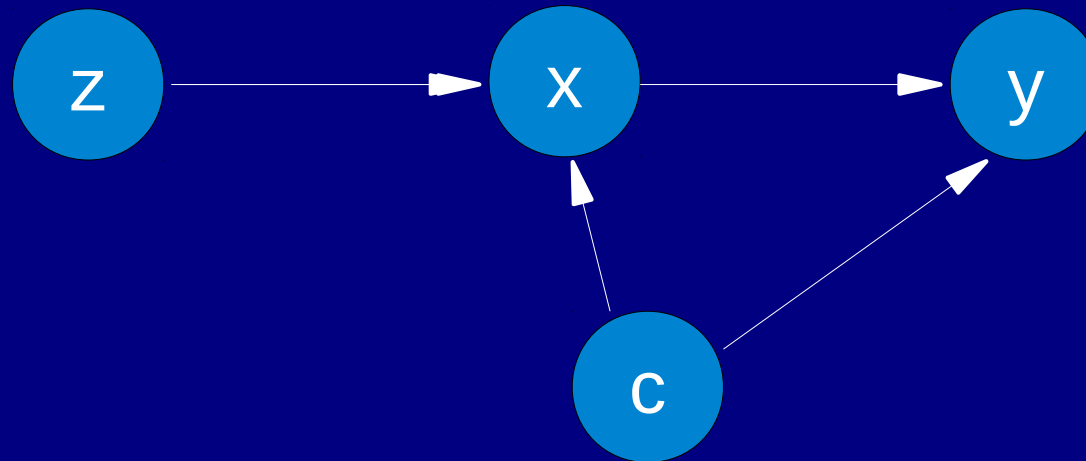
- Markov property:
 $p(x, y, z) = p(z) p(x | z) p(y | x)$
- $p(y | x, z) = p(y | x)$
- we say that y is conditionally independent of z given x
- but y and z are dependent

More on conditional dependencies in graphs

- What are the conditional dependencies in these graphs?



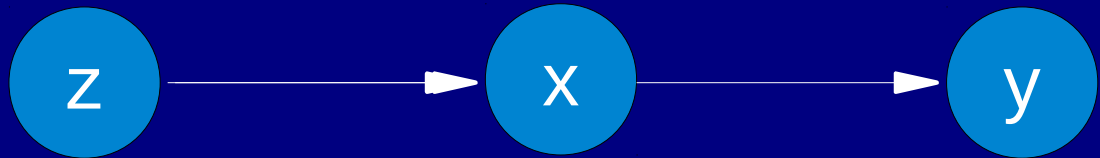
Graphical definition of a confounder



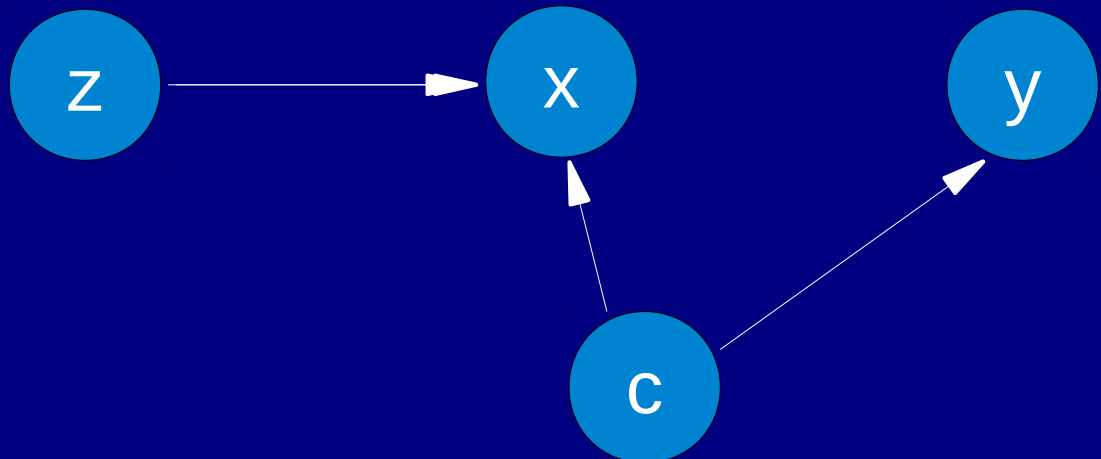
- Confounder of association between x and y is any variable on a pathway from which information flows to x and y
- information flow is defined by introducing a *do*-operator (equivalent to a latent instrumental variable z)
 - Information cannot flow backwards in time

Inferring causation from conditional independence in graphs

- Causal relationship: y depends on z



- Confounding: no information flow between z and y



Comparison of classical and Bayesian inference

- For inference about a model parameter, usually not much difference
 - Classical maximum likelihood estimates and confidence intervals are asymptotically equivalent to posterior modes and intervals
- For hypothesis testing (model comparison), Bayesian framework differs radically from classical methods
 - classical p -value evaluates null hypothesis ($\beta = \beta_0$) against a diffuse alternative
 - Bayesian test compares likelihood $P(\text{data} \mid \text{model})$ of alternative models

Comparison of Bayesian and classical inference about a parameter

- Classical approach: construct an “estimator” whose sampling distribution over repeated experiments can be calculated
 - asymptotic properties of maximum likelihood estimators
- Bayesian approach: compute the likelihood, and posterior density
 - with uninformative priors, likelihood and posterior are same
 - nothing is estimated, so we don't care about biased estimates

Bayesian hypothesis testing

- Prior odds: $P(H_1) / P(H_0)$
- Likelihood of hypothesis H given data Y : $P(Y | H)$
- Likelihood of a model with adjustable parameters is the *marginal likelihood* or *evidence*:
 - $P(Y | H) = \int P(Y | \theta) P(\theta | H) d\theta$
 - likelihood of parameters $P(Y | \theta)$, averaged over prior $P(\theta | H)$
 - Likelihood ratio (Bayes factor): $P(Y | H_1) / P(Y | H_0)$
- Posterior odds $P(H_1 | Y) / P(H_0 | Y) = \text{prior odds} \times \text{Bayes factor}$
 - $\log \text{posterior odds} = \log \text{prior odds} + \log \text{Bayes factor}$
- \log Bayes factor (lod score) is the weight of evidence favouring H_1 over H_0

Classical and Bayesian approaches to testing a null hypothesis against an alternative

- Classical p -value is calculated from difference between log-likelihood at the null and and log-likelihood at the maximum of the parameter.
 - evaluates null hypothesis against a diffuse alternative
- log Bayes factor is calculated from the difference between log-likelihoods of the null and the alternative hypotheses
 - alternative hypothesis must specify priors on parameters

Axiomatic basis for Bayesian hypothesis testing

- *Cox axioms*: degrees of belief must obey rules of probability if they satisfy simple criteria of logical consistency
- Bayesian inference uses rules of probability to revise degrees of belief given data
- Bayes theorem Implies **likelihood principle**
 - ratio of likelihoods contains all information in the data about the support for one hypothesis over another, or for one parameter value over another

Cox axioms

- Degrees of belief B can be ordered:
 - if $B(x) > B(y)$ and $B(y) > B(z)$, then $B(x) > B(z)$
- The degree of belief in a proposition x is related to the degree of belief in the negation of that proposition
$$B(x) = f [B(\text{not } x)]$$
- The degree of belief in the joint proposition (x and y) is related to the degree of belief in the conditional proposition (x given y) and the proposition y
 - $B(x \text{ and } y) = f [B(x | y), B(y)]$

Some other axiomatic bases for Bayesian inference

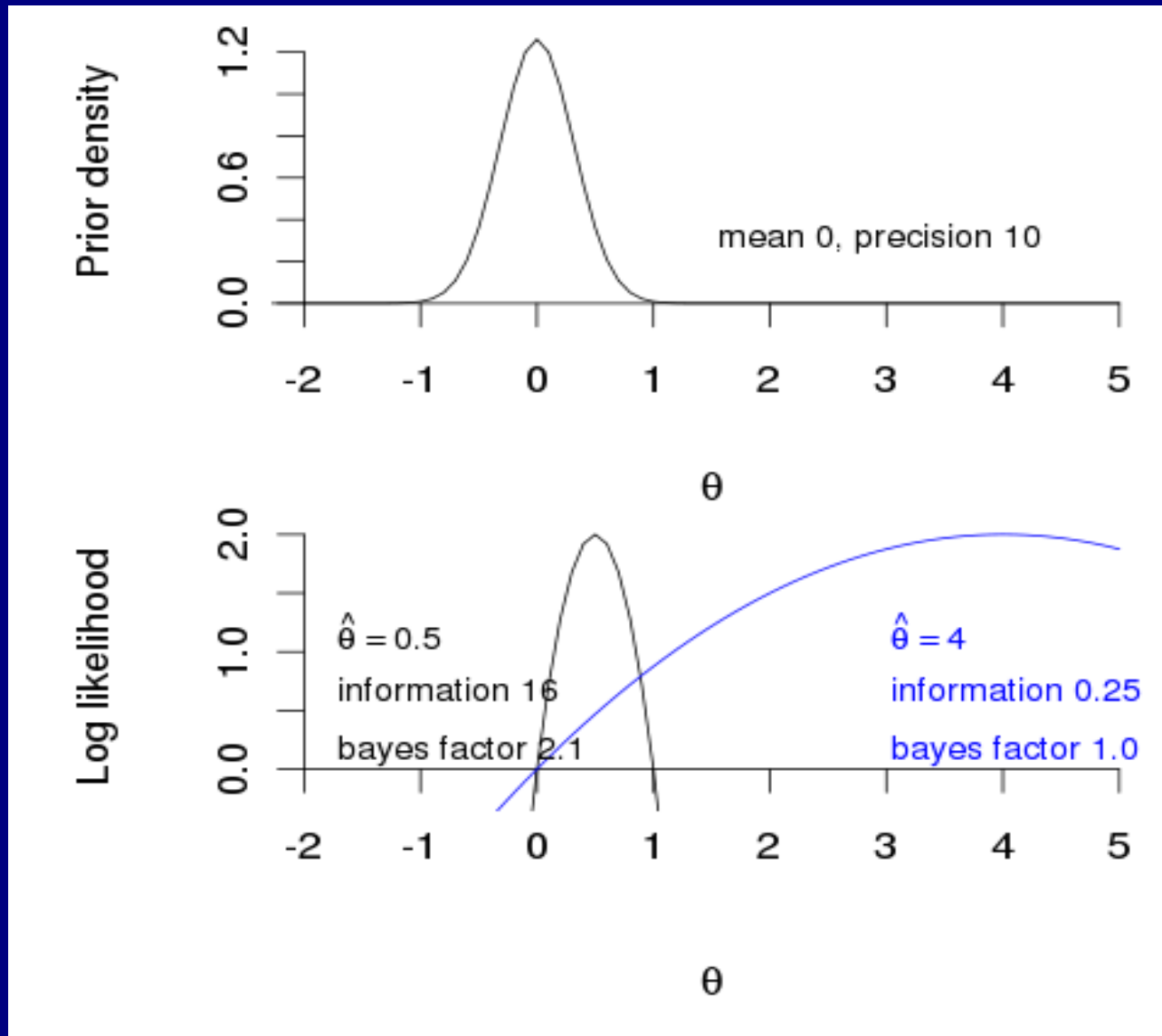
To update betting odds, strategy based on Bayesian inference will always win against any other strategy

- de Finetti's representation theorem
 - if 0-1 observations are exchangeable, it is as if they are independent samples from a Bernoulli distribution with a prior on the probability parameter
 - extensions: prior on circular errors
 - principle of maximum entropy

Interpretation of Bayesian hypothesis testing

- Evidence is quantified by ability of model (with priors on any adjustable parameters) to predict the data
 - Penalizes implausibly large effects in underpowered studies
 - Penalizes unnecessary complexity: model with highest marginal likelihood will be the simplest explanation that fits the data
- Validity does not depend on large sample approximations, or on having fewer variables than observations

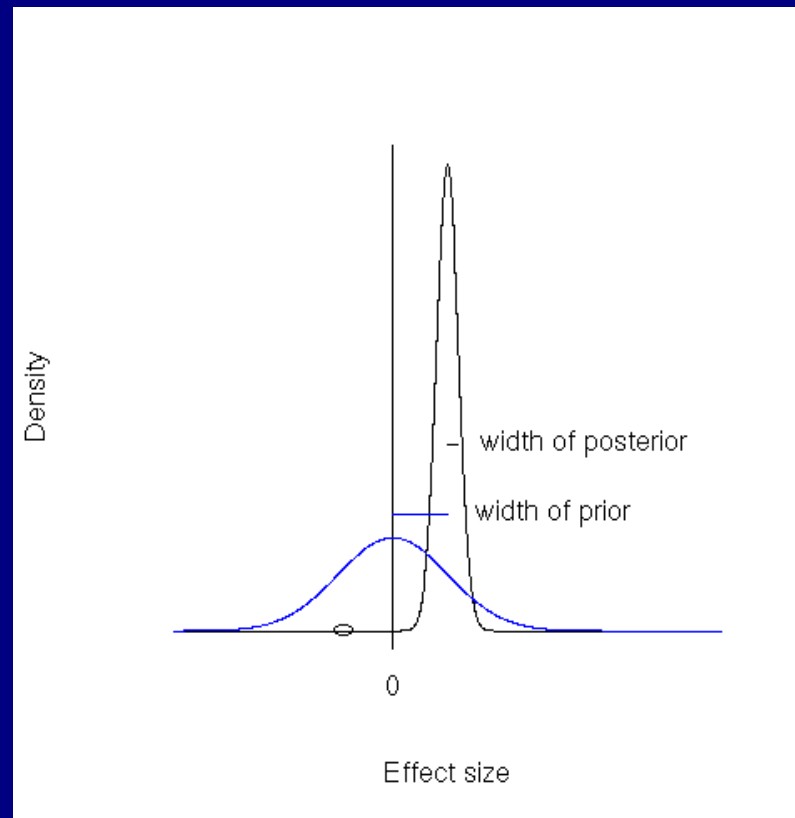
How Bayes factor penalizes implausibly large effects in underpowered studies



Bayesian interpretation of p -values

- Given a positive result in a diagnostic test
likelihood ratio = sensitivity / (1 – specificity)
- Significance test can be viewed as a diagnostic test:
 - threshold p -value = 1 – specificity
 - power for effect of plausible size = sensitivity
 - Likelihood ratio = power / threshold p -value
- p -values are misleading if study is underpowered to detect effects of plausible size

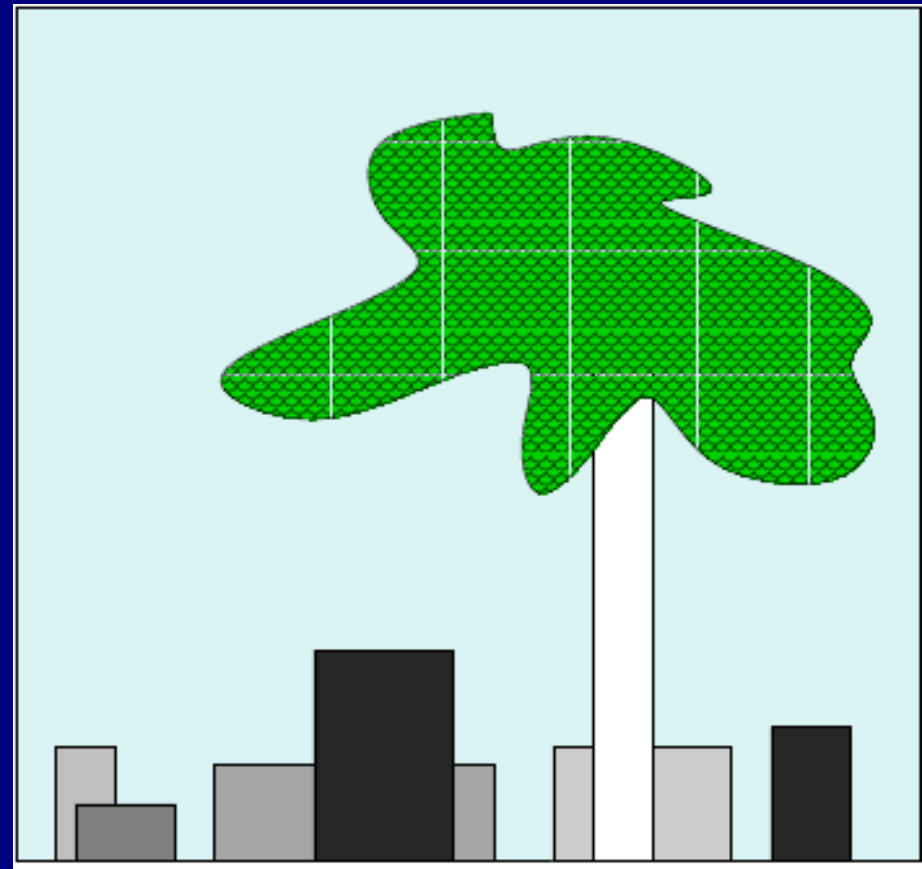
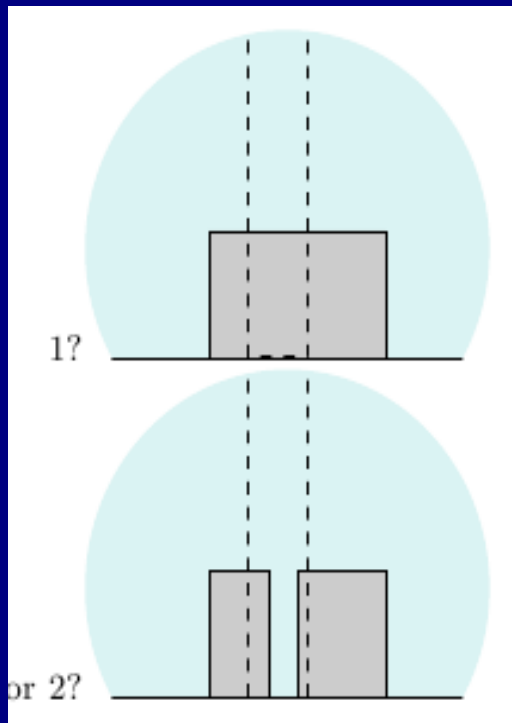
How marginal likelihood penalizes “complexity” (large prior hypothesis space)



- Occam factor = width of prior / width of posterior
- Likelihood of H_1 = best-fit likelihood x Occam factor

How Bayesian hypothesis testing favours the simplest explanation that fits the data: Mackay 2003

- How many boxes are behind the tree?



Two hypotheses: what is the likelihood ratio?

- H1: there is one box behind tree
 - 4 free parameters: 3 for coordinates of top and sides of box, 1 for colour of box
- H2: there are two boxes behind tree
 - 8 free parameters: 4 for each box
- Probability model for observations
 - x and y coordinates have 20 distinguishable values
 - tree is 3 units wide
 - box colour has 16 distinguishable values

Likelihood equivalence

- Likelihood equivalence of two hypotheses
 - given any setting of parameters of model H_1 , can find a setting of parameters of model H_0 such that both models have same likelihood $P(\text{data} | H)$ for all possible datasets
- Heckerman: priors should be set to ensure likelihood equivalence for models that have equivalent conditional dependencies
- Strict Bayesian argument: priors should describe beliefs.
 - models that are likelihood-equivalent may have different marginal likelihoods (evidence values)

Example: evidence even though causal effects are not identifiable (Mackay 2003)

	B=0	B=1
A=0	760	190
A=1	5	45

- 2 binary variables A and B
 - H_1 (A causes B): 3 parameters $P(A=1)$, $P(B | A=0)$, $P(B | A=1)$ have flat priors on 0, 1
 - H_2 (B causes A): 3 parameters $P(B=1)$, $P(A | B=0)$, $P(A | B=1)$ have flat priors on 0, 1
 - Bayes factor $P(H_1) / P(H_2)$ is 3.8

Exercise: calculate the likelihood ratio for hypothesis that A causes B, over hypothesis that B causes A

Given a uniform prior on the probability of success, the probability of r successes in n trials is

$$r! (n - r)! / (n + 1)!$$

– special case of the Beta-binomial likelihood

Why does the evidence favour H1 over H2?

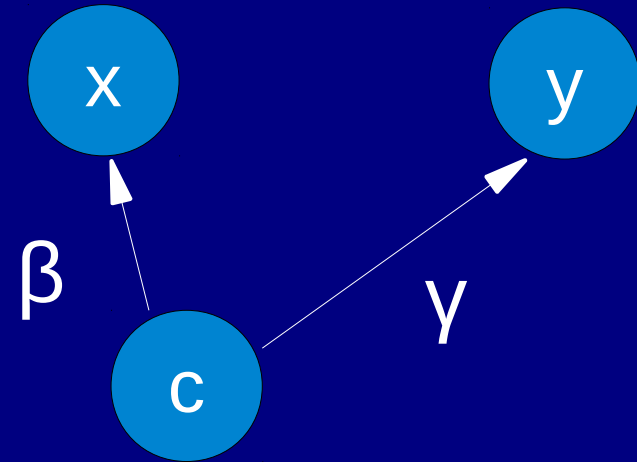
- Under H1, the maximum likelihood values of the probabilities are 0.95, 0.8, 0.1
- Under H2, the maximum likelihood values of the probabilities are 0.24, 0.008, 0.19

How does imposing a flat prior on probability of success encode extra information?

- Conjugate prior can be interpreted as “prior sample size”
- Principle of maximum entropy: given what you know, choose the prior that maximizes uncertainty (entropy)
 - otherwise you are encoding information that you don't have
- logistic regression model for dependence of B on A has maximal entropy given that model averages equate to data averages

Another example: strong associations favour causation over confounding

- Causal relationship: y depends on z
- Confounding: no information flow between z and y



For parameters to have same likelihood in both models, $\alpha = \beta \gamma$

Classical epidemiological approach to inferring causation from an exposure-disease association

- Measure all likely confounders: factors that are independently associated with outcome
- Test whether exposure-disease association persists after adjusting for these confounders
- Control of confounding is likely to fail with biomarkers because the likely confounders are unknown or difficult to measure
 - for instance raised cytokine levels predict age-related cognitive impairment – but are affected by underlying disease processes

Smoking and lung cancer debate in 1950s

- Classical statisticians' argument:
 - any inference of causation from observational data is unreliable
 - how do you know that all relevant confounders have been measured?
- Epidemiologists' argument
 - even without experimental confirmation, evidence from observational studies can strongly favour causation

Bradford Hill criteria: how to infer causation where classical criteria are not met

- Strength of association
- Temporal sequence
- Consistency
- Biological plausibility
- Coherence
- Specificity in the causes
- Dose-response relationship
- Experimental evidence
- Analogy

Bayesian interpretation of Bradford Hill criteria

- Strength of association: priors imply confounding effects are rarely strong
- Consistency and coherence: prior expectation that confounding effects will not be consistent
- Dose-response relationship: fits simple hypothesis of linear trend
- Specificity in causes: prior hypothesis space is small
- Biologic plausibility: high prior odds for causal relationship

Control of confounding: epidemiology faces its limits

- Standard methods for control of confounding in epidemiological studies are likely to fail if the exposure under study is:-
- A biomarker: e.g. an inflammatory marker
 - Association with outcome may be confounded by unknown metabolic/physiologic factors
- A health-seeking behaviour: e.g. use of vitamin E supplements, post-menopausal oestrogen
 - Association with outcome may be confounded by other health-seeking behaviours

Why does control of confounding fail for “endogenous” variables?

- Biomarkers:
 - confounders are unknown
 - Temporal sequence from exposure to outcome is difficult to establish: reverse causation is possible
- Behavioural factors
 - confounding is likely to be strong for a disease/outcome where risk can be modified by “lifestyle” factors
 - Measurement of exposure is often biased

Some failures in observational epidemiology

- Increased beta-carotene intake associated with lower risk of lung cancer (Peto et al. 1981, Willett 1990)
 - vitamin supplements advocated (Willett 2001)
 - RCTs of beta-carotene supplementation (1994) found no such evidence.
- Beta-carotene, vitamin E supplements, & hormone replacement therapy all predict lower risk of cardiovascular disease
 - all failed to be confirmed in RCTs.

Bayesian computation

- to learn parameters when you can specify the model
 - use Markov chain Monte Carlo (MCMC) simulation to sample the posterior density of model parameters
 - software (BUGS / JAGS) is available
- to learn which model is supported by the data
 - use approximate methods to compare model likelihoods
 - approximate the posterior by the mode (Laplace approximation), by a separable distribution (variational Bayes) or by gaussians (expectation propagation)

Bayesian approach to statistical models that have the form of a directed acyclic graph

- Specify *full probability model*: priors on all variables that have no “parents” in the graph
 - Can specify uninformative (“diffuse”) priors. With large samples and strong effects, priors have little influence on results
- Generate samples from the posterior distribution of unobserved variables using MCMC simulation
 - general-purpose software – BUGS, JAGS
 - Likelihood function of parameter θ obtained by weighting posterior samples of θ by inverse of prior density

Methods for evaluating the marginal likelihood of a model with adjustable parameters

- Averaging the likelihood over the prior distribution by quadrature is analogous to measuring volume of a lake by taking soundings
- surface may be high-dimensional, most of the volume may be in deep canyons
- Exact marginalization is possible only for special cases where the integrals are tractable
 - all gaussian or all discrete
 - no missing data or latent variables

Possible alternatives to evaluating marginal likelihoods: evaluate the ratio

To compare two models, we may be able to define a continuous parameter that includes both models as special cases

- e.g. define parameter θ as ratio of causal to crude (causal + confounding) effects
- Set any convenient prior on θ , then generate the posterior density
- Divide posterior by prior to get relative likelihood surface, and evaluate ratio of likelihoods at $\theta = 1$ and $\theta = 0$

Approximate methods for learning graphical models or computing marginal likelihood

- Laplace approximation
 - find the posterior mode of the parameters, then compute best-fit likelihood and the 2nd-derivative of the log posterior
 - Bayesian information criterion is a crude approximation to this
- Alpha-divergence methods
 - Variational Bayes: more accurate than Laplace approximation but not always tractable
 - Expectation-propagation: similar to variational Bayes, tractable but may fail to converge.

Software tools for Bayesian inference

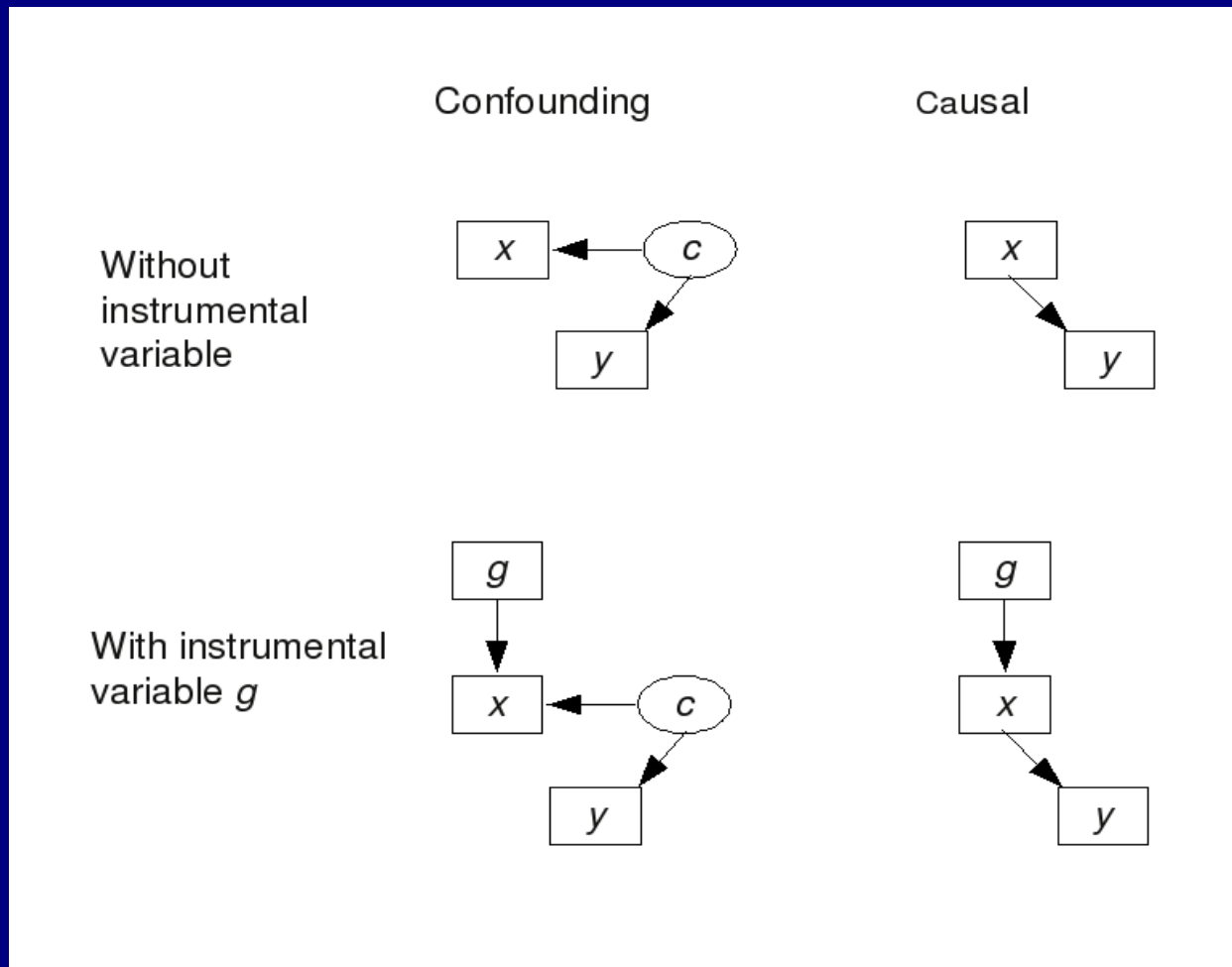
- BUGS and JAGS: specify model in a script, then use Markov chain Monte Carlo to generate posterior samples
- VIBES: specify your model in a graph, then use variational Bayes: teaching tool only.
- INFER.NET: has capabilities of both BUGS and VIBES. Specify model in C# script
- SPIV: can learn sparse instrumental variable model given data with many genotype and biomarker variables

Instrumental variable analysis

- Identify an “instrument” that perturbs the exposure of interest (usually a biomarker or behavioural factor)
- Assumptions:-
 - Effect of instrument on outcome is unconfounded
 - Any effect of instrument on outcome is mediated through the intermediate variable.
 - Effects of setting different levels of exposure are independent of the instrument

How instrumental variable analysis can distinguish causation from confounding

- exposure x , outcome y , unmeasured confounder c



Instrumental variable analysis in economics

- Economists want to infer the effects of “endogenous” (intermediate) variables that are likely to be confounded
- Example
 - age at leaving school is an “endogenous variable” that predicts lifetime earnings
 - variation in statutory school-leaving age can be used as an instrument
 - can estimate the causal effect of extra year's school on outcome

Instrumental variable analysis of clinical trials

- Standard “intention to treat” analysis ignores noncompliance
 - ok for hypothesis testing, but not for inferring size of treatment effect
- Can treat random allocation as the instrument, and treatment exposure as the intermediate variable
 - Allows size of treatment effect to be inferred with control for confounding by factors associated with non-compliance

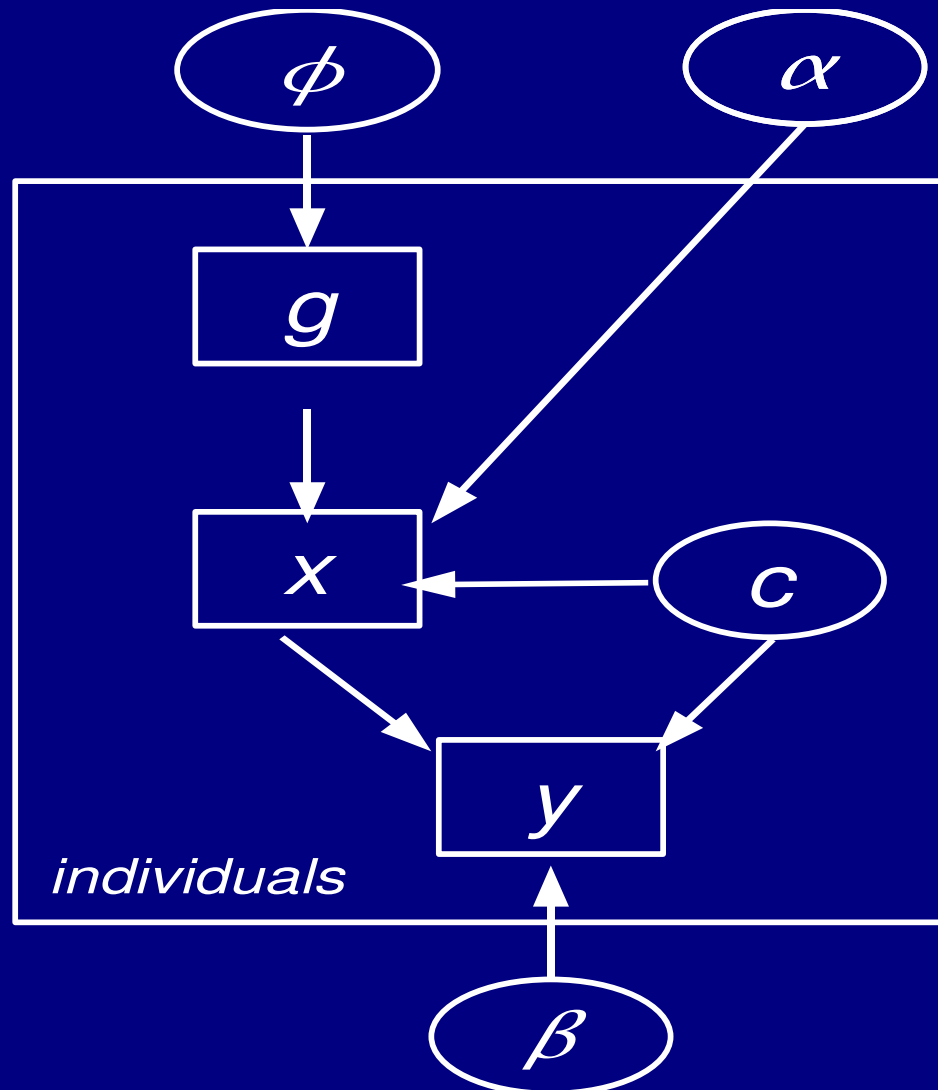
“Mendelian randomization”: instrumental variable analysis with genetic instruments

- Find one or more genes in which variation perturbs levels of the biomarker. Compare effects on outcome of
 - genetic perturbation of the biomarker
 - non-genetic variation of the biomarker
- Example:
 - raised plasma fibrinogen predicts cardiovascular disease
 - genotype in the beta-fibrinogen gene predicts fibrinogen levels
 - genotypic effects on fibrinogen levels do not predict cardiovascular disease

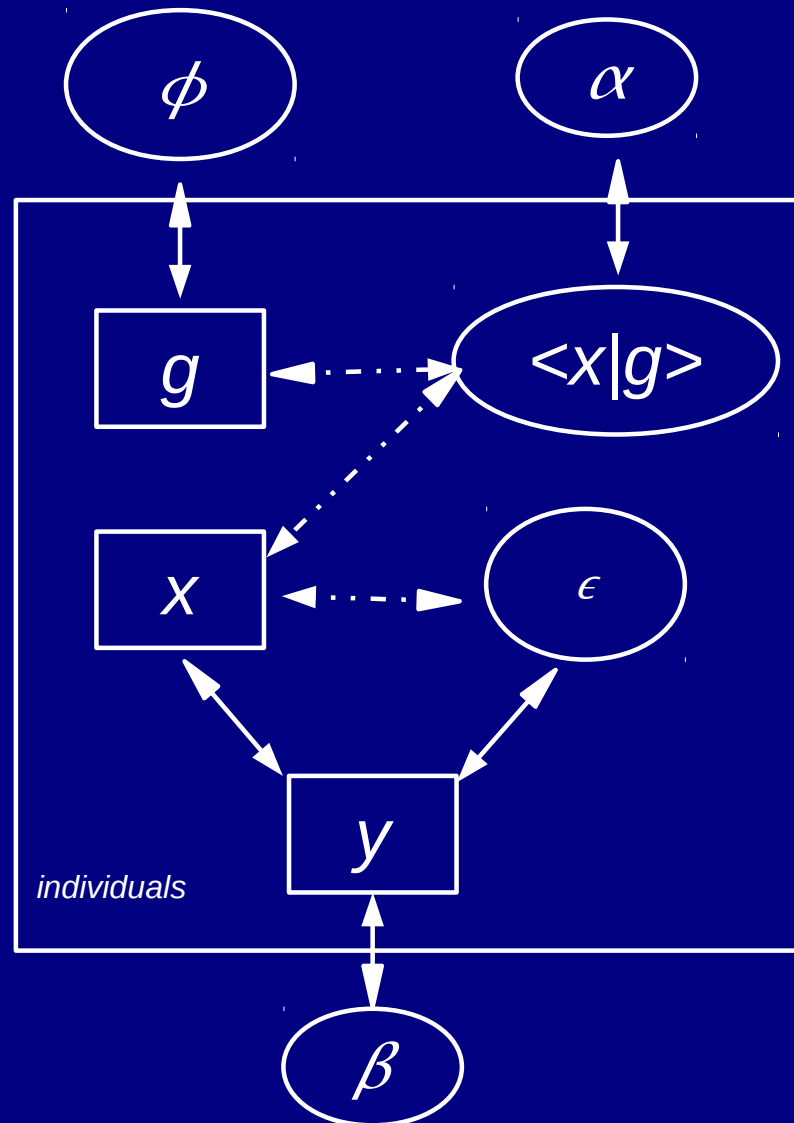
Assumptions underlying instrumental variable analysis with genetic instruments

- Effect of genotype on outcome is unconfounded
 - guaranteed by Mendel's laws, if population stratification is controlled
- Effect of genotype on outcome is mediated only through the intermediate phenotype (no *pleiotropy*)
- To be able to generalize: effects on outcome of different settings of the biomarker are independent of the instrument
 - no *developmental compensation / channelling*

Graphical model: genotype g as instrumental variable for effect of intermediate phenotype x on outcome y



Reparameterization: $x = \langle x|g \rangle + \epsilon$,
confounder c absorbed into random term ϵ



How instrumental variable model separates causal and confounding effects of biomarker on outcome

- Causal effect = effect on outcome of conditional expectation $\langle x | g \rangle$ of biomarker given genotype
- Causal + confounding effect = effect of residual deviation of biomarker from expectation given genotype

Inferring causal effect from relation of outcome to conditional expectation of intermediate phenotype given genotype

- Effect of genotype g on intermediate phenotype x

- $x = \alpha_0 + \alpha_g g + \epsilon = \langle x|g \rangle + \epsilon$

- Effect of intermediate phenotype x on outcome y

- $y = \beta_0 + \beta_x x + \beta_\epsilon \epsilon = \gamma_0 + \gamma_g g + \gamma_\epsilon \epsilon$

- where $\gamma_g = \beta_x \alpha_g$

- Causal effect parameter $\beta_x = \gamma_g / \alpha_g$

Inferring causal effect parameter: what priors are reasonable

- Causal effect parameter $\beta_x = \gamma_g / \alpha_g$
- but it is not appropriate to put independent priors on γ_g and α_g
- Classical ratio estimator is equivalent to Bayesian posterior mode with diffuse independent priors on γ_g and α_g
- Estimator behaves badly with weak instruments because priors are inappropriate

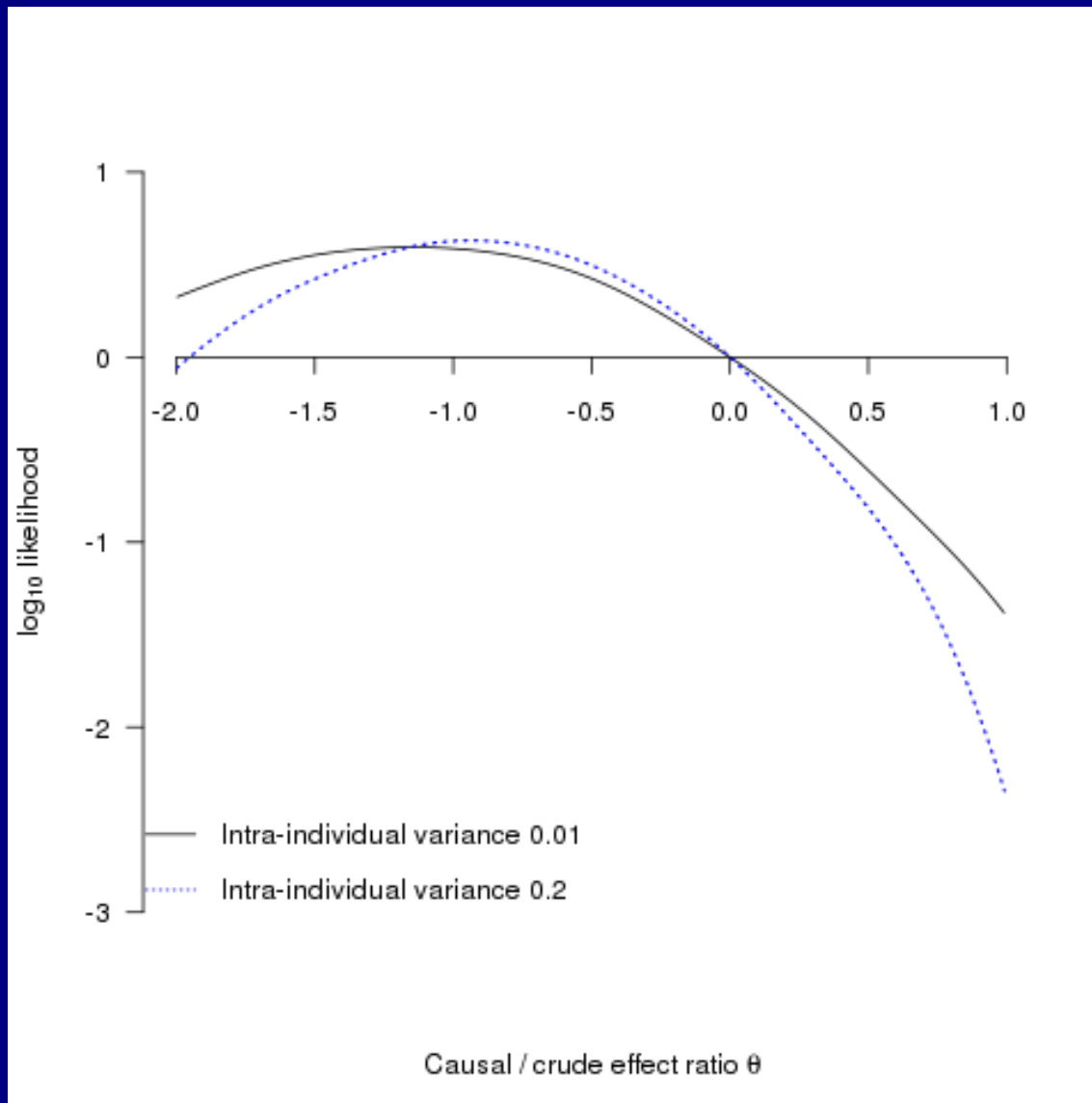
Hypothesis testing: define a parameter that spans causal and non-causal explanations

- θ is ratio of causal to crude (unadjusted) effect of intermediate phenotype x on outcome y
- General model: $y = \beta_0 + \gamma(\theta \langle x|g \rangle + \epsilon)$
 - angled brackets $\langle \rangle$ denote expectation
 - No causal effect ($\theta=0$): $y = \beta_0 + \gamma \epsilon$
 - All association of x with y is causal ($\theta=1$):
 $y = \beta_0 + \gamma x + \epsilon = \beta_0 + \gamma(\langle x|g \rangle + \epsilon)$
 - $\theta < 1$ would imply a causal effect opposite in direction to crude effect
- we can compute posterior density of θ

Example: *SLC2A9* genotype, urate, and metabolic syndrome in ORCADES

- Raised plasma urate levels are associated with metabolic syndrome
- ORCADES: 1017 adults examined for cardiovascular risk factors, urate
 - 706 typed for 5 SNPs in *SLC2A9*
- Regression slopes
 - urate on genotype: 0.22 ($p=2 \times 10^{-5}$)
 - metabolic syndrome on urate: 0.79 ($p=2 \times 10^{-13}$)
 - metabolic syndrome on genotype: -0.27 ($p=0.09$)

Log-likelihood of causal/crude effect ratio θ , with and without allowing for intra-individual variation of urate levels



Testing the assumptions of the model

- No residual population stratification:
 - Can test for stratification (EIGENSTRAT) using markers, estimate genetic background and adjust for it
- No pleiotropy
 - Can test this if multiple SNPs have been typed in the gene used as instrument
 - relative weights of SNPs should be same for effects on intermediate phenotype as for effects on biomarker
 - can construct a score test for this null hypothesis

Advantages of Bayesian approach to instrumental variable modelling

- Flexible: No need to assume linear relationships
 - classical instrumental variable methods are a special case of the Bayesian approach
- Does not rely on asymptotic large-sample properties
 - inference with weak instruments is valid
 - no need to construct “estimators”
- Evaluates weight of evidence (log Bayes factor) for causal over non-causal explanation of biomarker-disease association

Mendelian randomization studies need very large case-control collections

- Where:
 - N cases are required to detect effect of intermediate phenotype on outcome in a cohort study
 - Effect of genotype on intermediate phenotype is modest: 0.25 standard deviations for each extra copy of the trait-raising allele
 - allele frequency is 0.2
- $100 \times N$ cases are required to detect the corresponding effect of genotype on outcome in a case-control study

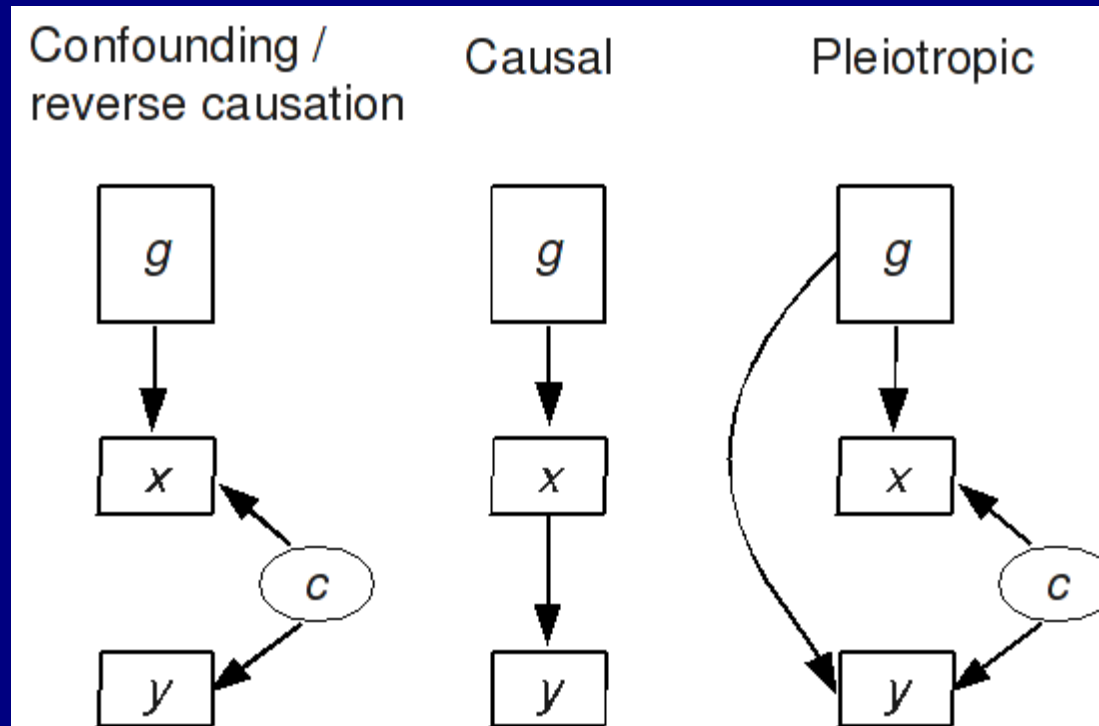
Extension to more general graphical models for effects of genotype, intermediate phenotype and environmental exposure on outcome

- Problem is to evaluate the posterior over models
 - requires computing marginal likelihood of each model
 - MCMC sampling (BUGS, JAGS) can generate the posterior distribution of parameters given a model, but (except in special cases) not the posterior over models
- Approximate methods (variational Bayes, expectation-propagation) are computationally tractable

More general approach: relax the assumption of no pleiotropy

- Instrumental variable argument assumes all effects of instrument on outcome are mediated through endogenous variable
- This assumption severely restricts ability to exploit genotype-biomarker associations for instrumental variable analysis
 - we rarely understand genetic effects well enough to assume no pleiotropic effects
- With multiple genetic instruments, and some model of genetic effects, possible in principle to relax this assumption

Can we infer causation where there may be pleiotropic effects of genotype on outcome?



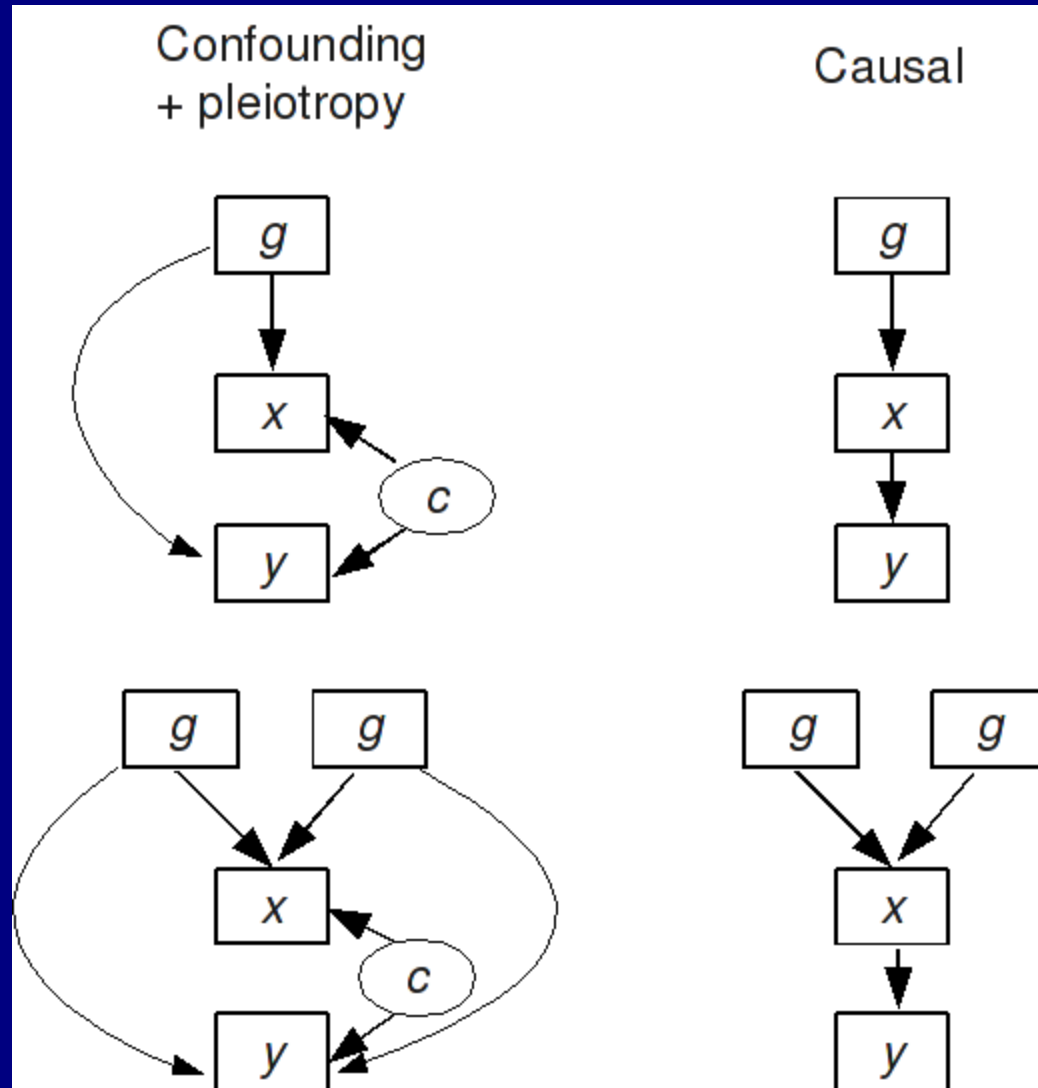
- “Mendelian randomization” argument assumes all genotypic effect on outcome is mediated through biomarker
- Schadt 2005: compares fit of these 3 models with a penalty for number of adjustable parameters)

Schadt et al (2005): “likelihood-based model causality selection” to detect causal effects

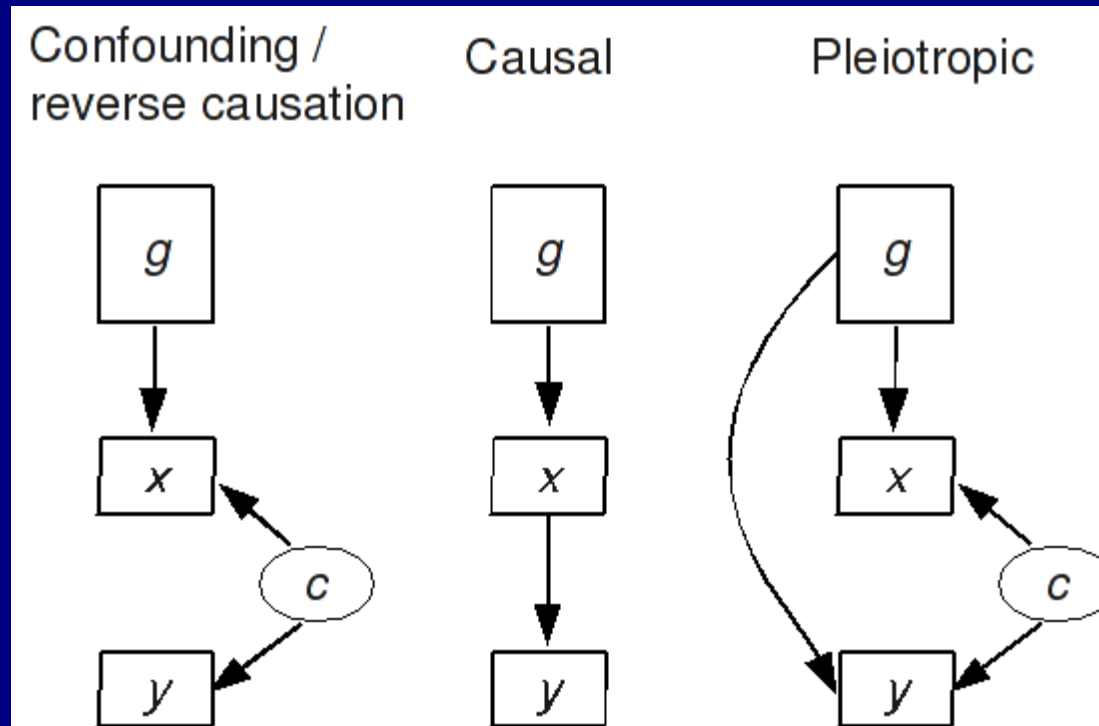
- Initial filtering step to select relevant genotypes and phenotypic biomarkers (gene transcript levels)
- For each possible genotype-transcript-trait triad, compares three possible models:-
 - confounding / reverse causation
 - causal
 - pleiotropic model with confounding
- Model choice is based on fit penalized by complexity (Akaike Information Criterion)

With multiple instruments, causal explanation can be distinguished from confounding + pleiotropy

- One instrument: 3 params versus 2
- Two instruments: 5 params versus 3



Schadt et al (2005) “likelihood-based causality model selection”



- Confounding and causal” models each have 2 effect parameters
- Pleiotropic model can fit the data perfectly (highest likelihood) but has 3 effect parameters (penalty for complexity)

Limitations of Schadt approach

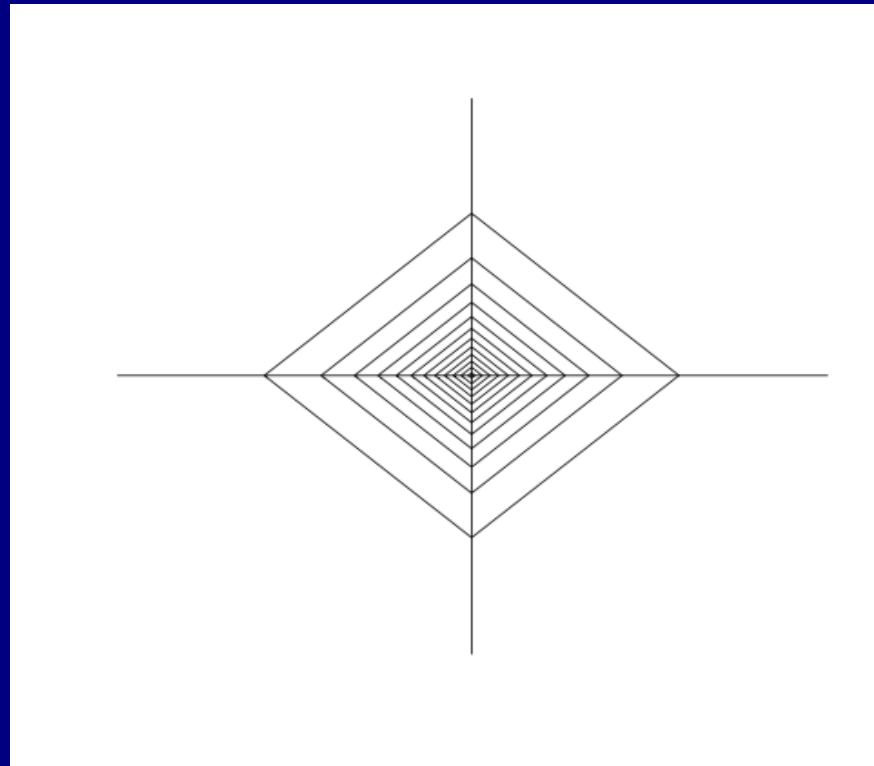
- Can only evaluate one genotype-biomarker-outcome triad at a time
- Does not allow for noisy measurements
- Choice between models with and without pleiotropy depends upon an arbitrary penalty for complexity
 - AIC is inappropriate for hypothesis testing
 - Formal hypothesis test based on marginal likelihood would depend critically on
 - priors on model parameters
 - assumptions about measurement noise

Sparse Bayesian linear models

- Sparse prior: prior probabilities are highest for models in which most effect parameters are zero
- Can be set up for *automatic relevance determination*: effects not supported (models with low likelihood) are “pruned” from the model as it is fitted to the data
- we don't have to specify the extent of sparseness: can learn a “sparseness” parameter from the data

How double exponential prior encodes sparsity

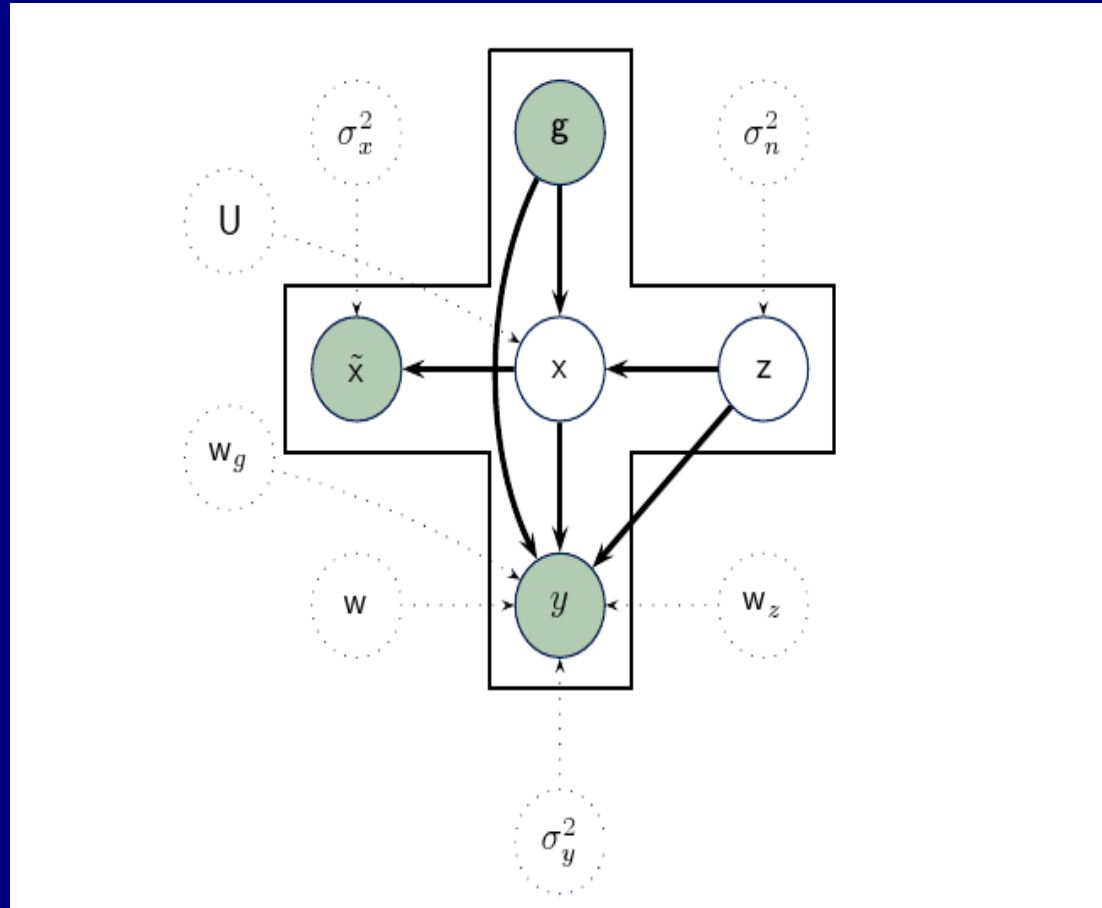
- Density varies inversely with sum of absolute values of effect parameters
- posterior is log-concave: can use EM algorithm to fit model



Sparse Bayesian instrumental variable analysis

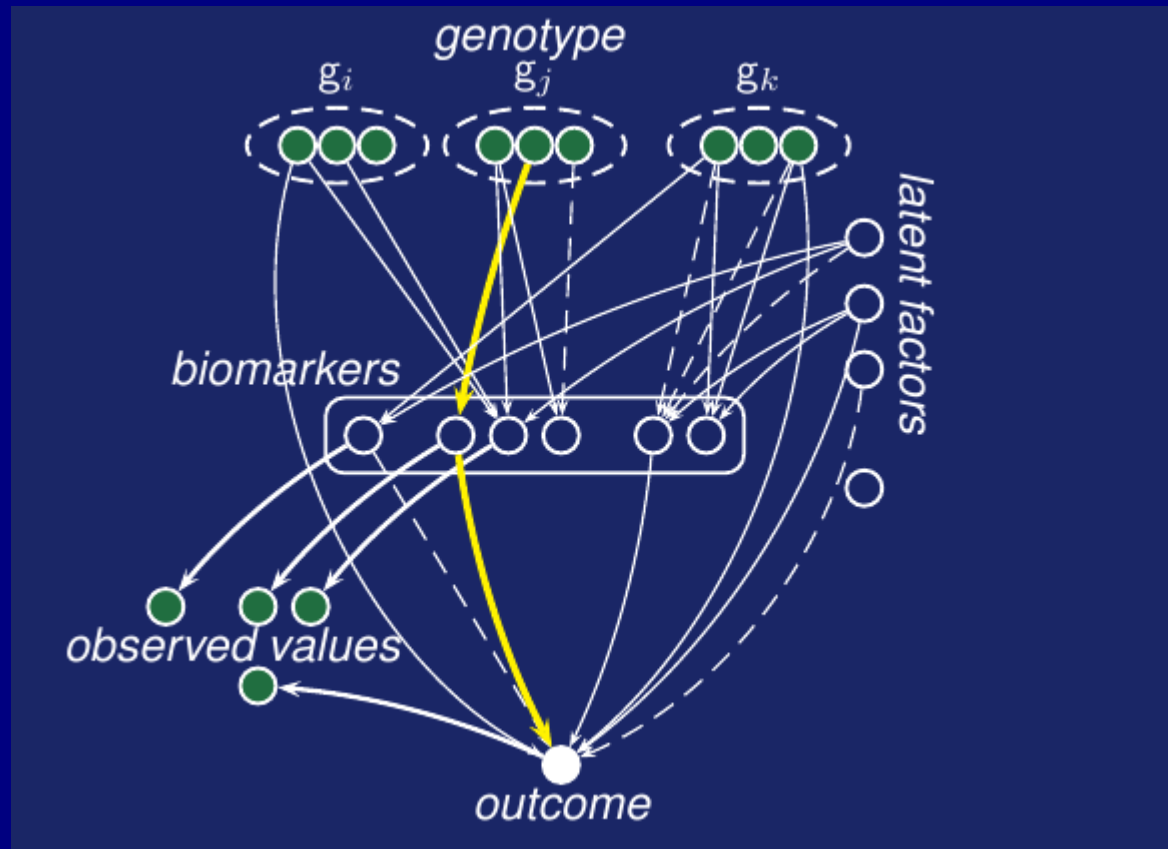
- Initial filtering step to select biomarkers associated with outcome and genotypes associated with biomarkers
- Specify model with all possible genotype-biomarker-outcome links, including pleiotropic effects of genotype on outcome
- Laplace (double exponential) priors on all effect parameters: encodes sparseness
- Convex optimization algorithm iterates to posterior mode
 - *automatic relevance determination*: effects not supported by the data are pruned from the model

General model with multiple instruments, pleiotropy and unobserved confounding



- Observed variables: filled circles
- Unobserved variables: clear circles

Full model for genotype-biomarker-outcome associations

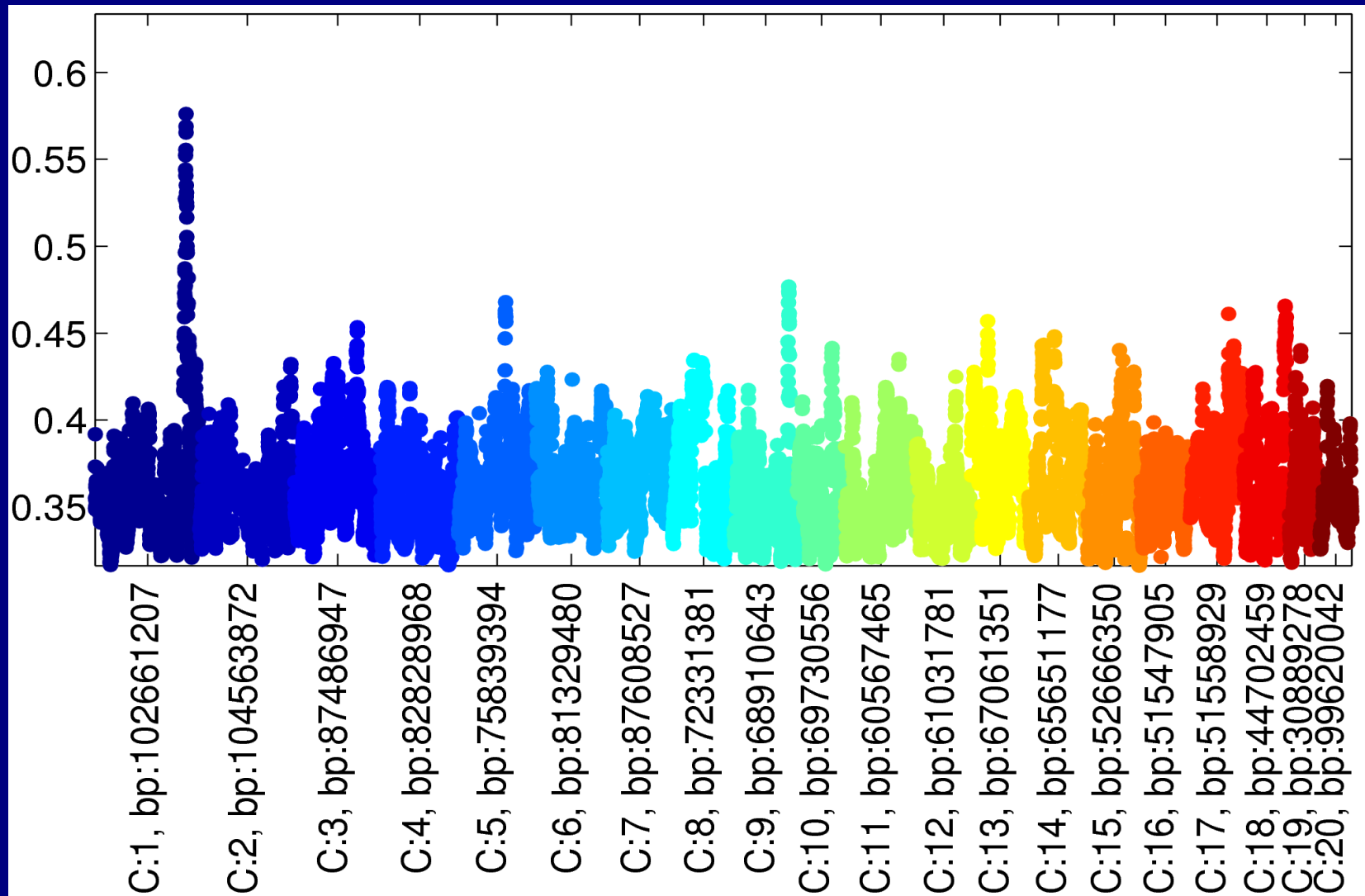


- Confounders encoded as latent factors that may be common to multiple biomarkers
- Observed values of biomarkers are noisy

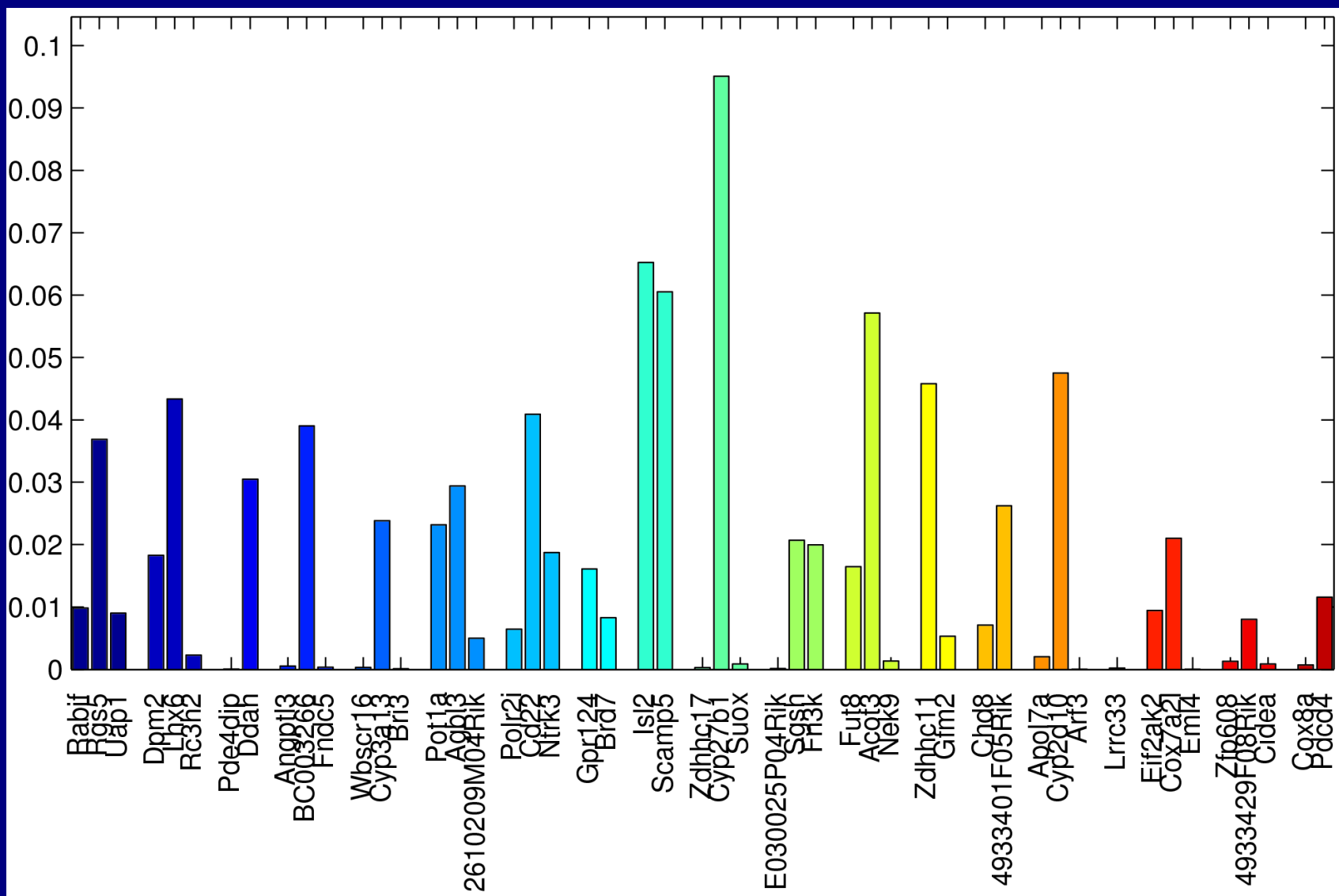
Liver transcript levels as biomarkers for plasma HDL cholesterol

- 260 mice from a “heterogeneous stock” formed by crossing 8 inbred strains over ~ 100 generations
- Measured
 - genotypes (ancestry from founder strains) at marker loci
 - gene expression in liver (47000 transcripts) retaining only those informative for genotype
 - plasma lipids
- Sparse Bayesian instrumental variable analysis with HDL level as outcome

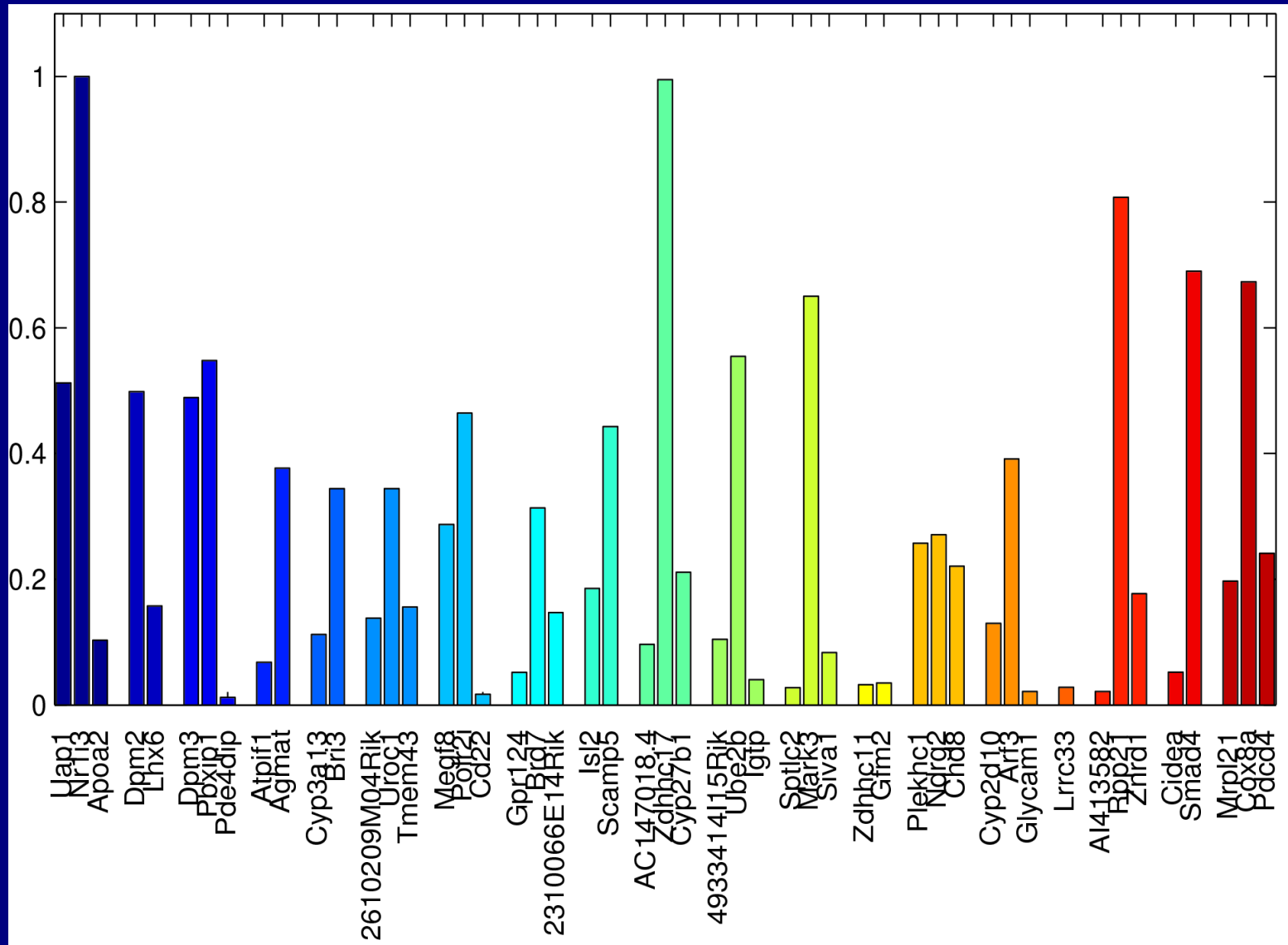
Mutual information between HDL cholesterol and ancestry at genome-wide marker loci in HS mice



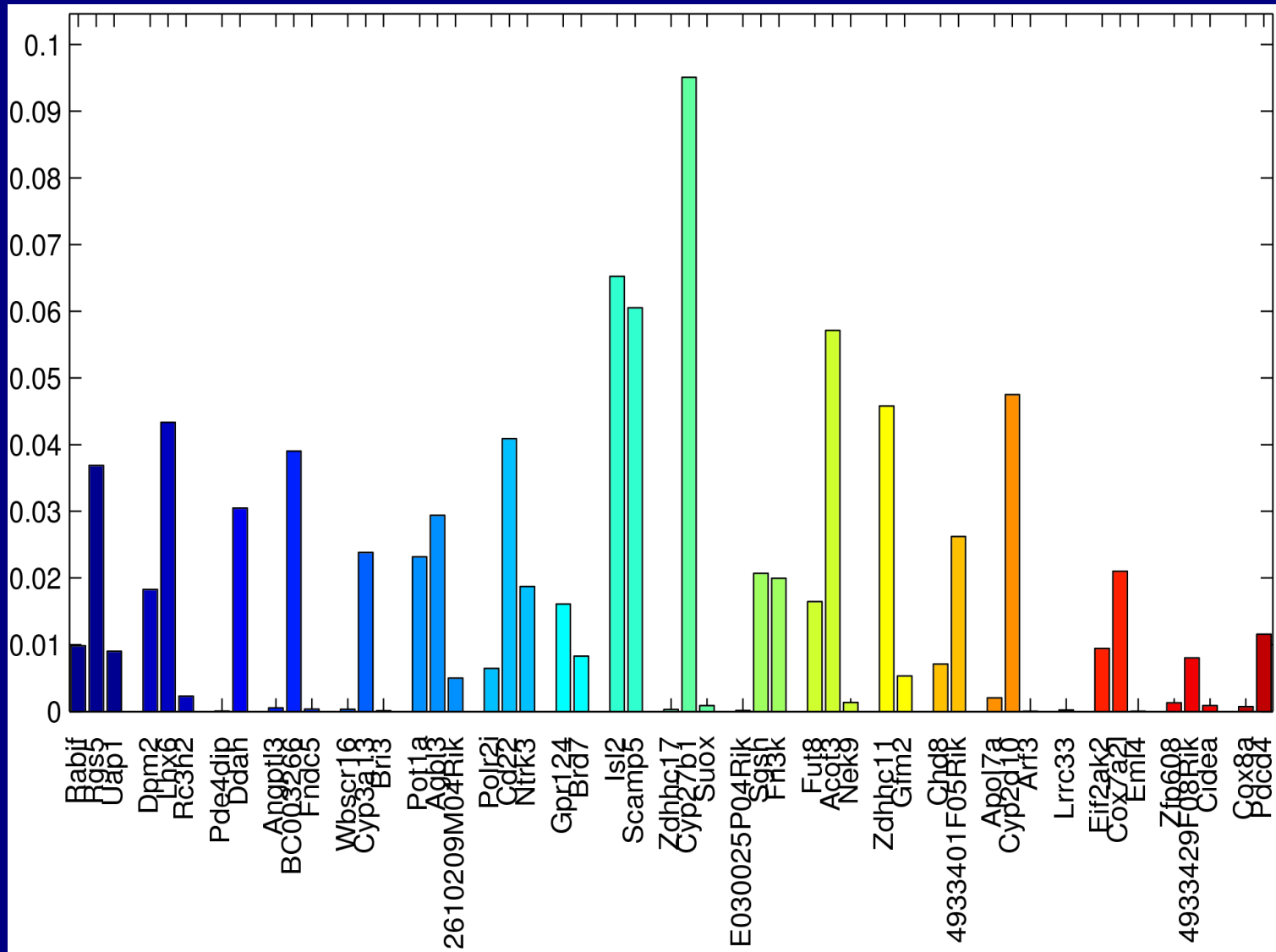
Using ancestry at marker loci to infer causal effects of transcript levels in liver on HDL cholesterol in HS mice



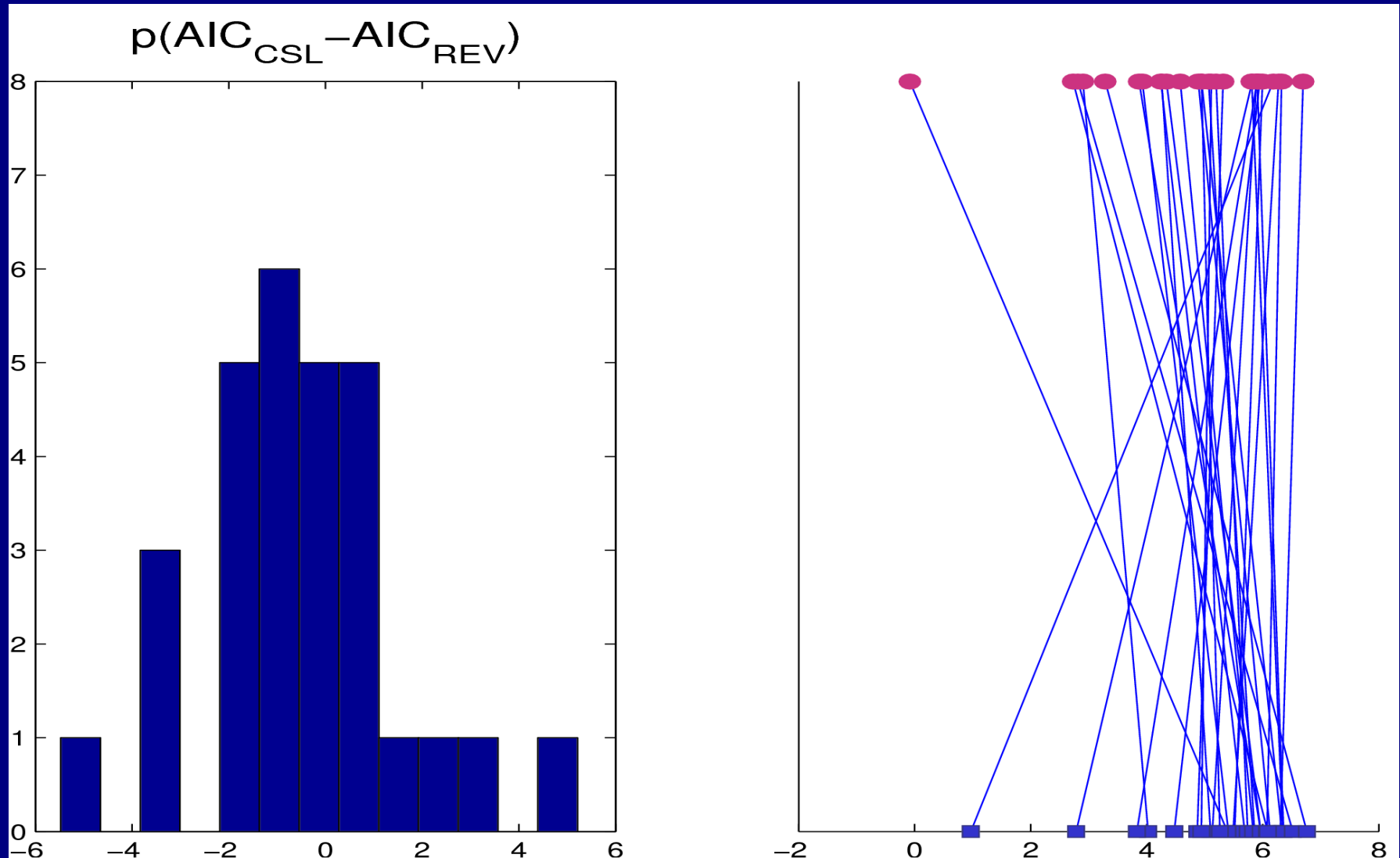
Fine mapping: transcripts that contain information about genotypic effects on HDL



Inferred causal effects of transcript levels on HDL



Comparison with Schadt's "likelihood-based model causality" test applied to CYP27B1



Summary (1) : causal models

- Causal relations can be represented by information flow in graphical models
- Classical framework allows causal effects to be inferred from conditional independence relationships under certain conditions (randomized instrument, or no unobserved confounders)
- In principle, genotypic variation can be exploited to infer causal effects of phenotypic biomarkers on outcomes

Summary (2): Bayesian hypothesis testing

- Bayesian hypothesis testing is based on comparing (marginal) likelihoods
- In the Bayesian framework, inference about causality does not necessarily depend upon conditional independence in graphs
- Bayesian inference automatically penalizes unnecessarily complex explanations

Summary (3) “Mendelian randomization”

- Mendelian randomization applies the classical instrumental variable argument to inferring causal relations between biomarkers and outcomes
 - Depends critically upon the “no pleiotropy” assumption
 - Bayesian inference using MCMC overcomes limitations of classical “estimation” methods
- Application is restricted to specific genes and biomarkers where the biology is well understood

Summary (4): learning genotype-biomarker outcome relationships where pleiotropic effects are not excluded

- With multiple genotypic instruments, it is possible to relax the assumption of no pleiotropy
- Causal explanations will be favoured if they explain the data equally well as more complex hypotheses involving pleiotropy and confounding
- In principle this can be applied to investigate all genotypes and biomarkers simultaneously

Summary (5): sparse Bayesian instrumental variable analysis

- Unified computational framework for studying genotypic and biomarker effects on outcome
- Sparse priors allow automatic relevance determination: instead of searching over all possible models
- Can infer causal relationships between biomarkers and outcomes relationships
- Software is under rapid development.