

Redes Bayesianas: Cómo se aprenden? (II)

Aritz Pérez¹ Borja Calvo²

Basque Center for Applied Mathematics

UPV/EHU

Donostia, Febrero de 2015

Bibliografía

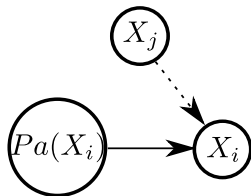
Koller09: D. Koller y M. Friedman (2009). Probabilistic Graphical Models. MIT Press.

Castillo97: E. Castillo, J.M. Gutiérrez, y A.S. Hadi (1997). Sistemas Expertos y Modelos de Redes Probabilísticas. Academia de Ingeniería.

Aproximación cualitativa

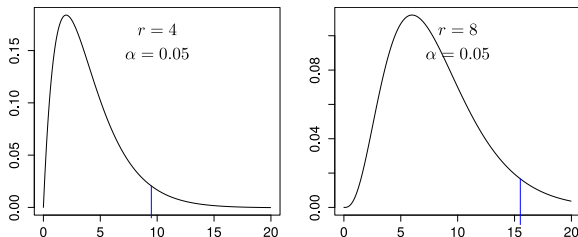
- Modelar las **dependencias** mas importantes de p
- Emplea el **test de la independencia condicionada** para decidir cuando resulta beneficioso modelar una dependencia
- Determinar el manto de Markov

Test de la independencia condicionada



- Pasamos del factor $p(X_i | \mathbf{pa}(X_i))$ al factor $p(X_i | \mathbf{pa}(X_i), x_j)$
- Es equivalente a **rechazar** la independencia $i(X_i; X_j | \mathbf{Pa}(X_i))$
- Test de independencia condicional con $\hat{l}(X_i; X_j | \mathbf{Pa}(X_i))$

Control de la complejidad



- Los **grados de libertad** aumentan **exponencialmente** con el número de padres $, (r_i - 1)(r_j - 1)(r_{Pa_i})$
- El **test** se vuelve **más exigente**

Alternativas al test de independencia

- Emplear $1 - pval$ o $\alpha - pval$ como funciones de evaluación
- Test de la verosimilitud:
 - **Generalización** del test de la independencia
 - Comparar **un grafo con otro** con arcos adicionales
 - Grados de libertad de la χ^2 es la **diferencia de parámetros**

Aprendizaje exacto de redes Bayesianas

- **Programación dinámica**: Subproblemas superpuestos
- **Integer linear programming**: Función lineal sujeta a restricciones lineales
- **Weighted max-sat**: Codificar como problema de satisfabilidad
- Muy **costosos**: Aplicables a datos con **pocas variables**

Heurísticos de búsqueda: Algoritmo voraz

- Búsqueda local
- Vecindario: grafos con un **arco** adicional
- **Eficiente**
- **Optimos locales**
- Variantes: otros vecindarios, búsquedas tabu, rollout,...

Heurísticos de búsqueda: Basados en poblaciones

- **Individio** representa un grafo
- Algoritmos **genéticos**
- Algoritmos de **estimación de distribuciones**
- Basados en **colonias**