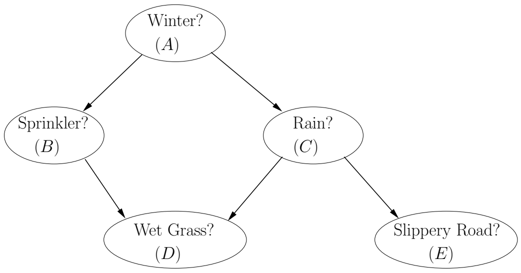
Probabilistic Models: Spring 2014
Scoring Functions Example

We are given the following Bayesian network $\mathcal{N}$.



We are also given the following dataset $\mathcal{D}$.

A	B	C	D	E	Count
T	F	T	T	T	20
T	F	F	F	F	15
F	T	F	T	T	10
F	F	T	T	T	15
F	F	F	F	F	5
T	T	F	T	F	2

For all of the calculations, we need the counts, n_{ijk}. They are as follows.

A	PA_A	i	j	k	n_{ijk}
T	$\emptyset$	1	1	1	37
F	$\emptyset$	1	1	2	30

B	PA_B (A)	i	j	k	n_{ijk}
T	T	2	1	1	2
F	T	2	1	2	35
T	F	2	2	1	10
F	F	2	2	2	20

C	PA_C (A)	i	j	k	n_{ijk}
T	T	3	1	1	20
F	T	3	1	2	17
T	F	3	2	1	15
F	F	3	2	2	15

E	PA_E (C)	i	j	k	n_{ijk}
T	T	5	1	1	35
F	T	5	1	2	0
T	F	5	2	1	10
F	F	5	2	2	22

D	PA_D (BC)	i	j	k	n_{ijk}
T	TT	4	1	1	0
F	TT	4	1	2	0
T	TF	4	2	1	12
F	TF	4	2	2	0
T	FT	4	3	1	35
F	FT	4	3	2	0
T	FF	4	4	1	0
F	FF	4	4	2	20

1. Calculate the MDL score for the network
   For example, we can calculate the MDL score of *Sprinkler* (B) given its parent set, which is $\{Winter\ (A)\}$.

$$\begin{aligned}
MDL(B:\mathcal{D},\mathcal{N}) &= -\left\{\sum_j^{q_i}\sum_k^{r_i} N_{ijk}\log\frac{N_{ijk}}{N_{ij}}\right\} + \frac{\log N}{2}\cdot(r_i-1)\cdot q_i \\
&= -\left\{\sum_j^{2}\sum_k^{2} N_{2jk}\log\frac{N_{2jk}}{N_{2j}}\right\} + \frac{\log 67}{2}\cdot(2-1)\cdot 2 \\
&= -\left\{N_{211}\log\frac{N_{211}}{N_{21}} + N_{212}\log\frac{N_{212}}{N_{21}} + N_{221}\log\frac{N_{221}}{N_{22}} + N_{222}\log\frac{N_{222}}{N_{22}}\right\} + \frac{\log 67}{2}\cdot 2 \\
&= -\left\{2\log\frac{2}{37} + 35\log\frac{35}{37} + 10\log\frac{10}{30} + 20\log\frac{10}{30}\right\} + \log 67 \\
&\approx 31.08
\end{aligned}$$

   Note that log is calculated as the natural logarithm, as is common in practice.

   Using similar calculations, we see that the scores of the other variables are as follows. The score of the entire network is their sum.

X	$MDL(X:\mathcal{D},\mathcal{N})$
A	48.17
B	31.08
C	50.52
D	8.41
E	24.08

2. Calculate the BDeu score for the network with ESS=0.1

   Again, we will calculate the score for B.

$$
\begin{aligned}
BDeu(B:\mathcal{D},\mathcal{N}) &= \sum_{j}^{q_i} \log\Gamma(\frac{\alpha}{q_i}) - \log\Gamma(\frac{\alpha}{q_i} + N_{ij}) + \\
&\qquad \sum_{k}^{r_i} \log\Gamma(\frac{\alpha}{r_i \cdot q_i} + N_{ijk}) - \log\Gamma(\frac{\alpha}{r_i \cdot q_i}) \\
&= \sum_{j}^{2} \log\Gamma(\frac{0.1}{2}) - \log\Gamma(\frac{0.1}{2} + N_{2j}) + \\
&\qquad \sum_{k}^{2} \log\Gamma(\frac{0.1}{4} + N_{2jk}) - \log\Gamma(\frac{0.1}{4}) \\
&= \log\Gamma(0.05) - \log\Gamma(0.05 + N_{21}) + \\
&\qquad \log\Gamma(0.025 + N_{211}) - \log\Gamma(0.025) + \log\Gamma(0.025 + N_{211}) - \log\Gamma(0.025) + \\
&\qquad \log\Gamma(0.05) - \log\Gamma(0.05 + N_{22}) + \\
&\qquad \log\Gamma(0.025 + N_{221}) - \log\Gamma(0.025) + \log\Gamma(0.025 + N_{221}) - \log\Gamma(0.025) \\
&= \log\Gamma(0.05) - \log\Gamma(0.05 + 37) + \\
&\qquad \log\Gamma(0.025 + 2) - \log\Gamma(0.025) + \log\Gamma(0.025 + 35) - \log\Gamma(0.025) + \\
&\qquad \log\Gamma(0.05) - \log\Gamma(0.05 + 30) + \\
&\qquad \log\Gamma(0.025 + 10) - \log\Gamma(0.025) + \log\Gamma(0.025 + 20) - \log\Gamma(0.025) \\
&\approx -35.14
\end{aligned}
$$

Note that $\log\Gamma(\cdot)$ was evaluated using the built-in *lgamma* function in C++.

Similarly, the scores of the other variables are as follows.

X	$BDeu(X:\mathcal{D},0.1,\mathcal{N})$
A	-50.31
B	-35.14
C	-55.42
D	-2.21
E	-25.13

3. Calculate the BDeu for the network with ESS=100

   By replacing α with 100 instead of 0.1, we find the following scores.

X	$BDeu(X:\mathcal{D},100,\mathcal{N})$
A	-46.55
B	-39.83
C	-46.90
D	-29.53
E	-38.55

Useful Equations

Minimum description length

$$MDL(\mathcal{N}:\mathcal{D}) = -\sum_i^n \left\{ \sum_j^{q_i} \sum_k^{r_i} N_{ijk} \log \frac{N_{ijk}}{N_{ij}} \right\} + \frac{\log_2 N}{2} \cdot (r_i - 1) \cdot q_i$$

$$MDL(\mathcal{N}:\mathcal{D}) = -\sum_i^n \ell(X_i|PA_i) + \frac{\log_2 N}{2} \cdot (r_i - 1) \cdot q_i$$

Bayesian Dirichlet with likelihood equivalence and uninformative priors

$$\begin{aligned}
P(\mathcal{D},\mathcal{N}) &= P(\mathcal{N})P(\mathcal{D}|\,\mathcal{N}) && \text{Rewrite using chain rule}\\
&= P(\mathcal{N}) \prod_i^n \prod_j^{q_i} \frac{\Gamma(\alpha_{ij})}{\Gamma(\alpha_{ij}+N_{ij})} \prod_k^{r_i} \frac{\Gamma(\alpha_{ijk}+N_{ijk})}{\Gamma(\alpha_{ijk})} && \text{Substitute probability of data}\\
&\propto \prod_i^n \prod_j^{q_i} \frac{\Gamma(\alpha_{ij})}{\Gamma(\alpha_{ij}+N_{ij})} \prod_k^{r_i} \frac{\Gamma(\alpha_{ijk}+N_{ijk})}{\Gamma(\alpha_{ijk})} && \text{Assume a uniform structure prior}\\
&\propto \prod_i^n \prod_j^{q_i} \frac{\Gamma(\frac{\alpha}{q_i})}{\Gamma(\frac{\alpha}{q_i}+N_{ij})} \prod_k^{r_i} \frac{\Gamma(\frac{\alpha}{r_i \cdot q_i}+N_{ijk})}{\Gamma(\frac{\alpha}{r_i \cdot q_i})} && \text{Replace the } \alpha\text{s}\\
BDeu(\mathcal{N}:\mathcal{D},\alpha) &= \sum_i^n \left\{ \sum_j^{q_i} \log \frac{\Gamma(\frac{\alpha}{q_i})}{\Gamma(\frac{\alpha}{q_i}+N_{ij})} + \sum_k^{r_i} \log \frac{\Gamma(\frac{\alpha}{r_i \cdot q_i}+N_{ijk})}{\Gamma(\frac{\alpha}{r_i \cdot q_i})} \right\} && \text{Work in log-space}\\
BDeu(\mathcal{N}:\mathcal{D},\alpha) &= \sum_i^n \left\{ \sum_j^{q_i} \log\Gamma(\frac{\alpha}{q_i}) - \log\Gamma(\frac{\alpha}{q_i}+N_{ij}) + \right. && \text{Remove divisions}\\
&\qquad \left. \sum_k^{r_i} \log\Gamma(\frac{\alpha}{r_i \cdot q_i}+N_{ijk}) - \log\Gamma(\frac{\alpha}{r_i \cdot q_i}) \right\}
\end{aligned}$$