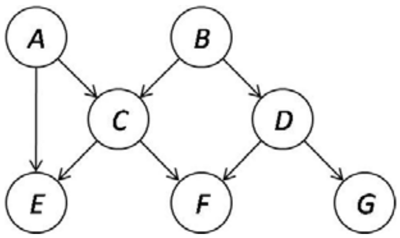
Exercises lecture 5

1. From the following Bayesian network, compute $p(A|E,F)$ in terms of known conditional probabilities:



2. a) Prove the Laplace approximation:

$$p(D|M) \approx p\left(D|M,\hat{\theta}_{\text{MAP}}\right) p\left(\hat{\theta}_{\text{MAP}}|M\right) \frac{1}{\sqrt{\det \frac{\mathbf{H}}{2\pi}}}$$

with $\mathbf{H} = -\nabla\nabla \log p(\theta|D,M)\big|_{\theta=\hat{\theta}_{\text{MAP}}}$.

b) What is **H** when the posterior over θ is a multivariate Gaussian with mean $\hat{\theta}_{\text{MAP}}$ and covariance matrix $\mathbf{\Sigma}$?

3. The following exercise from MacKay, *Information theory, inference, and learning algorithms*

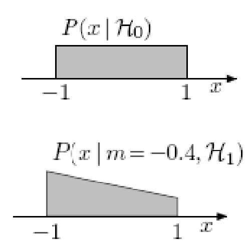
Exercise 28.1.[3] Random variables x come independently from a probability distribution $P(x)$. According to model $\mathcal{H}_0$, $P(x)$ is a uniform distribution

$$P(x|\mathcal{H}_0) = \frac{1}{2} \qquad x \in (-1,1). \tag{28.20}$$

According to model $\mathcal{H}_1$, $P(x)$ is a nonuniform distribution with an unknown parameter $m \in (-1,1)$:

$$P(x|m,\mathcal{H}_1) = \frac{1}{2}(1+mx) \qquad x \in (-1,1). \tag{28.21}$$

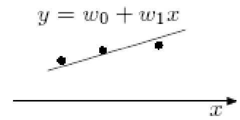
Given the data $D = \{0.3, 0.5, 0.7, 0.8, 0.9\}$, what is the evidence for $\mathcal{H}_0$ and $\mathcal{H}_1$?



4. The following exercise from the same book:

Exercise 28.2.[3] Datapoints (x, t) are believed to come from a straight line. The experimenter chooses x, and t is Gaussian-distributed about

$$y = w_0 + w_1 x \tag{28.22}$$



with variance σ_ν^2. According to model $\mathcal{H}_1$, the straight line is horizontal, so $w_1 = 0$. According to model $\mathcal{H}_2$, w_1 is a parameter with prior distribution Normal$(0, 1)$. Both models assign a prior distribution Normal$(0, 1)$ to w_0. Given the data set $D = \{(-8, 8), (-2, 10), (6, 11)\}$, and assuming the noise level is $\sigma_\nu = 1$, what is the evidence for each model?