31.[illegible]1 Advanced Machine Learning

a.) we have $p(\vec{x}) = \frac{1}{Z(\beta)} \exp(-\beta \vec{E}(\beta))$

$\ln(1/p(\vec{x})) = \ln(Z(\beta)) + \ln(\exp(\beta E(\beta)))$

$= \ln(Z(\beta)) + \beta E(\beta)$

$\rightarrow S = \sum p(\vec{x}) (\ln(1/p(\vec{x})))$

$= \ln(Z(\beta)) + \beta \vec{E}(\beta) \left(\sum_x p(\vec{x}) \right)$

↑ This is normalized, and thus equal to 1

$\Rightarrow S = \ln(Z(\beta)) + \beta \vec{E}(\beta)$ □

b.) $-\frac{\partial F}{\partial T} = \frac{-\partial F}{\partial \beta} \frac{\partial \beta}{\partial T} = \frac{-\partial F}{\partial \beta} \left(\frac{\partial}{\partial T} \left(\frac{1}{T} \right) \right) = \frac{1}{T^2} \frac{\partial F}{\partial \beta}$

$= \beta^2 \frac{\partial F}{\partial \beta}$ where $F = -1/\beta \ln(Z(\beta))$

$\Rightarrow \frac{\partial F}{\partial \beta} = \frac{1}{\beta^2} \ln(Z(\beta)) - \frac{1}{\beta} \frac{\partial}{\partial \beta} \ln(Z(\beta))$

$\rightarrow -\partial F / \partial T = \ln(Z(\beta)) + -\beta \frac{\partial}{\partial \beta} \ln(Z(\beta))$

(eq 31.8 Mck) $= \ln(Z(\beta)) + \beta \vec{E}(\beta) = \boxed{-\partial F / \partial T = S}$

Dirren van Vlijmen
Matthijs Neefelings
Olivier Brahma