14.26

$p(x^t)T(x^{t-1}|x^t)=p(x^{t-1})T(x^t|x^{t-1})$

[解析]：假设变量x所在的空间有n个状态$(s_1,s_2,..,s_n)$, 定义在该空间上的一个转移矩阵$T(n\times n)$如果满足一定的条件则该马尔可夫过程存在一个稳态分布π, 使得

$$\pi T=\pi \tag{1}$$

其中, π是一个是一个n维向量，代表$s_1,s_2,..,s_n$对应的概率. 反过来, 如果我们希望采样得到符合某个分布π的一系列变量$x_1,x_2,..,x_t$, 应当采用哪一个转移矩阵$T(n\times n)$呢？

事实上，转移矩阵只需要满足马尔可夫细致平稳条件

$$\pi(i)T(i,j)=\pi(j)T(j,i) \tag{2}$$

即公式14.26，这里采用的符号与西瓜书略有区别以便于理解. 证明如下

$$\pi T(j)=\sum_i \pi(i)T(i,j)=\sum_i \pi(j)T(j,i)=\pi(j) \tag{3}$$

假设采样得到的序列为$x_1,x_2,..,x_{t-1},x_t$，则可以使用MH算法来使得x_{t-1}(假设为状态s_i)转移到x_t(假设为状态s_j)的概率满足式(2).

14.28

$A(x^*|x^{t-1})=\min\left(1,\frac{p(x^*)Q(x^{t-1}|x^*)}{p(x^{t-1})Q(x^*|x^{t-1})}\right)$

[推导]：这个公式其实是拒绝采样的一个trick，因为基于式14.27只需要

$$\begin{aligned}A(x^*|x^{t-1})&=p(x^*)Q(x^{t-1}|x^*)\\A(x^{t-1}|x^*)&=p(x^{t-1})Q(x^*|x^{t-1})\end{aligned} \tag{4}$$

即可满足式14.26，但是实际上等号右边的数值可能比较小，比如各为0.1和0.2，那么好不容易才到的样本只有百分之十几得到利用，所以不妨将接受率设为0.5和1，则细致平稳分布条件依然满足，样本利用率大大提高, 所以可以将(4)改进为

$$\begin{aligned}A(x^*|x^{t-1})&=\frac{p(x^*)Q(x^{t-1}|x^*)}{norm}\\A(x^{t-1}|x^*)&=\frac{p(x^{t-1})Q(x^*|x^{t-1})}{norm}\end{aligned} \tag{5}$$

其中

$$norm=\max\left(p(x^{t-1})Q(x^*|x^{t-1}),p(x^*)Q(x^{t-1}|x^*)\right) \tag{6}$$

即教材的14.28.

14.32

$\ln p(x)=\mathcal{L}(q)+\mathrm{KL}(q\parallel p)$

[推导]：根据条件概率公式$p(x,z)=p(z|x)*p(x)$，可以得到$p(x)=\frac{p(x,z)}{p(z|x)}$

然后两边同时作用$\ln$函数，可得$\ln p(x) = \ln\frac{p(x,z)}{p(z|x)}$ (1)

因为$q(z)$是概率密度函数，所以$1 = \int q(z)dz$

等式两边同时乘以$\ln p(x)$，因为$\ln p(x)$是不关于变量z的函数，所以$\ln p(x)$可以拿进积分里面，得到$\ln p(x) = \int q(z)\ln p(x)dz$

$$\begin{aligned}
\ln p(x) &= \int q(z)\ln p(x) \\
&= \int q(z)\ln\frac{p(x,z)}{p(z|x)} \qquad (\text{带入公式}(1)) \\
&= \int q(z)\ln\left\{\frac{p(x,z)}{q(z)}\cdot\frac{q(z)}{p(z|x)}\right\} \\
&= \int q(z)\left(\ln\frac{p(x,z)}{q(z)} - \ln\frac{p(z|x)}{q(z)}\right) \\
&= \int q(z)\ln\left\{\frac{p(x,z)}{q(z)}\right\} - \int q(z)\ln\frac{p(z|x)}{q(z)} \\
&= \mathcal{L}(q) + \mathrm{KL}(q \parallel p) \qquad (\text{根据}\mathcal{L}\text{和}\mathrm{KL}\text{的定义})
\end{aligned}$$

14.36

$$\begin{aligned}
\mathcal{L}(q) &= \int \prod_i q_i \left\{\ln p(\mathbf{x},\mathbf{z}) - \sum_i \ln q_i\right\} d\mathbf{z} \\
&= \int q_j \left\{\int p(x,z)\prod_{i\neq j} q_i d\mathbf{z_i}\right\} d\mathbf{z_j} - \int q_j \ln q_j d\mathbf{z_j} + \mathrm{const} \\
&= \int q_j \ln\tilde{p}(\mathbf{x},\mathbf{z_j}) d\mathbf{z_j} - \int q_j \ln q_j d\mathbf{z_j} + \mathrm{const}
\end{aligned}$$

[推导]：

$$\mathcal{L}(q) = \int \prod_i q_i \left\{\ln p(\mathbf{x},\mathbf{z}) - \sum_i \ln q_i\right\} d\mathbf{z} = \int \prod_i q_i \ln p(\mathbf{x},\mathbf{z}) d\mathbf{z} - \int \prod_i q_i \sum_i \ln q_i d\mathbf{z}$$

公式可以看做两个积分相减，我们先来看左边积分$\int \prod_i q_i \ln p(\mathbf{x},\mathbf{z}) d\mathbf{z}$的推导。

$$\begin{aligned}
\int \prod_i q_i \ln p(\mathbf{x},\mathbf{z}) d\mathbf{z} &= \int q_j \prod_{i\neq j} q_i \ln p(\mathbf{x},\mathbf{z}) d\mathbf{z} \\
&= \int q_j \left\{\int \ln p(\mathbf{x},\mathbf{z}) \prod_{i\neq j} q_i d\mathbf{z_i}\right\} d\mathbf{z_j} \qquad (\text{先对}\mathbf{z_j}\text{求积分，再对}\mathbf{z_i}\text{求积分})
\end{aligned}$$

这个就是教材中的14.36左边的积分部分。

我们现在看下右边积分的推导$\int \prod_i q_i \sum_i \ln q_i d\mathbf{z}$的推导。

在此之前我们看下$\int \prod_i q_i \ln q_k d\mathbf{z}$的计算

$$\begin{aligned}
\int \prod_i q_i \ln q_k d\mathbf{z} &= \int q_{i'} \prod_{i\neq i'} q_i \ln q_k d\mathbf{z} \qquad (\text{选取一个变量}q_{i'}, i' \neq k) \\
&= \int q_{i'} \left\{\int \prod_{i\neq i'} q_i \ln q_k d\mathbf{z_i}\right\} d\mathbf{z_{i'}}
\end{aligned}$$

$\left\{\int\prod_{i\neq i'} q_i \ln q_k d\mathbf{z_i}\right\}$部分与变量$q_{i'}$无关，所以可以拿到积分外面。又因为$\int q_{i'}\, d\mathbf{z_{i'}}\ = 1$，所以

$$\begin{aligned}\int\prod_i q_i \ln q_k d\mathbf{z} &= \int\prod_{i\neq i'} q_i \ln q_k d\mathbf{z_i} \\ &= \int q_k \ln q_k d\mathbf{z_k} \qquad \text{（所有}k\text{以外的变量都可以通过上面的方式消除）}\end{aligned}$$

有了这个结论，我们再来看公式

$$\begin{aligned}\int\prod_i q_i \sum_i \ln q_i d\mathbf{z} &= \int\prod_i q_i \ln q_j d\mathbf{z} + \sum_{k\neq j}\int\prod_i q_i \ln q_k d\mathbf{z} \\ &= \int q_j \ln q_j d\mathbf{z_j} + \sum_{z\neq j}\int q_k \ln q_k d\mathbf{z_k} \qquad \text{（根据上面结论）} \\ &= \int q_j \ln q_j d\mathbf{z_j} + \mathrm{const} \qquad \text{（这里我们关心的是}q_j\text{，其他变量可以视为const）}\end{aligned}$$

这个就是14.36右边的积分部分。

14.40

$$q_j^*(\mathbf{z}_j) = \frac{\exp\left(\mathbb{E}_{i\neq j}[\ln(p(\mathbf{x},\mathbf{z}))]\right)}{\int\exp\left(\mathbb{E}_{i\neq j}[\ln(p(\mathbf{x},\mathbf{z}))]\right)\mathrm{d}\mathbf{z}_j}$$

[推导]：由14.39去对数并积分

$$\begin{aligned}\int q_j^*(\mathbf{z}_j)\mathrm{d}\mathbf{z}_j &= \int\exp\left(\mathbb{E}_{i\neq j}[\ln(p(\mathbf{x},\mathbf{z}))]\right)\cdot\exp(const)\,\mathrm{d}\mathbf{z}_j \\ &= \exp(const)\int\exp\left(\mathbb{E}_{i\neq j}[\ln(p(\mathbf{x},\mathbf{z}))]\right)\,\mathrm{d}\mathbf{z}_j \\ &= 1\end{aligned} \tag{7}$$

所以

$$\exp(const) = \frac{1}{\int\exp\left(\mathbb{E}_{i\neq j}[\ln(p(\mathbf{x},\mathbf{z}))]\right)\,\mathrm{d}\mathbf{z}_j} \tag{8}$$

$$\begin{aligned}q_j^*(\mathbf{z}_j) &= \exp\left(\mathbb{E}_{i\neq j}[\ln(p(\mathbf{x},\mathbf{z}))]\right)\cdot\exp(const)\,\mathrm{d}\mathbf{z}_j \\ &= \frac{\exp\left(\mathbb{E}_{i\neq j}[\ln(p(\mathbf{x},\mathbf{z}))]\right)}{\int\exp\left(\mathbb{E}_{i\neq j}[\ln(p(\mathbf{x},\mathbf{z}))]\right)\mathrm{d}\mathbf{z}_j}\end{aligned} \tag{9}$$