KL divergence of factored distributions

Notation:

- $p(x)$ and $q(x)$ are two different distributions of scalar random variable x
- $p(y)$ and $q(y)$ are two different distributions of scalar random variable y

Derivation:

$$
\begin{aligned}
& D_{\mathrm{KL}}(p(x)p(y)||q(x)q(y)) \\
= \; & \mathbb{E}_{p(x)p(y)}[\log(p(x)p(y)) - \log(q(x)q(y))] \\
= \; & \sum_{x,y} p(x)p(y)[\log(p(x)p(y)) - \log(q(x)q(y))] \\
= \; & \sum_{x,y} p(x)p(y)[\log p(x) - \log q(x) + \log p(y) - \log q(y)] \\
= \; & \sum_{x,y} p(x)p(y)[\log p(x) - \log q(x)] + \sum_{x,y} p(x)p(y)[\log p(y) - \log q(y)] \\
= \; & \sum_{x} \sum_{y} p(x)p(y)[\log p(x) - \log q(x)] + \sum_{x} \sum_{y} p(x)p(y)[\log p(y) - \log q(y)] \\
= \; & \sum_{x} p(x)[\log p(x) - \log q(x)] + \sum_{y} p(y)[\log p(y) - \log q(y)] \\
= \; & D_{\mathrm{KL}}(p(x)||q(x)) + D_{\mathrm{KL}}(p(y)||q(y))
\end{aligned}
$$