My solutions to

Deep Learning: Foundations and Concepts

Dario Miro Konopatzki

3 Standard Distributions

3.1

$$\begin{aligned}
\sum p_{X;\mu} &= \sum_{x\in\{0,1\}} \mu^x(1-\mu)^{1-x} \\
&= \mu^0(1-\mu)^{(1-0)} + \mu^1(1-\mu)^{1-1} \\
&= 1\cdot(1-\mu) + \mu\cdot 1 \\
&= 1-\mu+\mu \\
&= 1
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{x\in\{0,1\}} x\mu^x(1-\mu)^{1-x} \\
&= 0 + 1\cdot\mu^1(1-\mu)^{1-1} \\
&= \mu
\end{aligned}$$

$$\begin{aligned}
\mathbb{V}[X] &= \mathbb{E}\left[x^2\right] - \mathbb{E}[x]^2 \\
&= \sum_{x\in\{0,1\}} x^2\mu^x(1-\mu)^{1-x} - \mu^2 \\
&= 0 + 1^2\cdot\mu^1(1-\mu)^{1-1} - \mu^2 \\
&= \mu-\mu^2 \\
&= \mu(1-\mu)
\end{aligned}$$

$$\begin{aligned}
\mathrm{H}[x] &= \mathbb{E}\left[-\log_2 p\right] \\
&= \sum_{x\in\{0,1\}} -p(x)\log_2 p(x)
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{x\in\{0,1\}} -\mu^x(1-\mu)^{1-x}\log_2\left(\mu^x(1-\mu)^{1-x}\right) \\
&= \sum_{x\in\{0,1\}} -\mu^x(1-\mu)^{1-x}\left(x\log_2\mu + (1-x)\log_2(1-\mu)\right) \\
&= -\mu^0(1-\mu)^{1-0}\left(0\cdot\log_2\mu + (1-0)\log_2(1-\mu)\right) \\
&\quad -\mu^1(1-\mu)^{1-1}\left(1\cdot\log_2\mu + (1-1)\log_2(1-\mu)\right) \\
&= (1-\mu)\log_2(1-\mu) - \mu\log_2\mu
\end{aligned}
$$

3.2

$$
\begin{aligned}
\sum p_{X;\mu} &= \sum_{\{-1,1\}} \left(\frac{1-\mu}{2}\right)^{\frac{1-x}{2}} \left(\frac{1+\mu}{2}\right)^{\frac{1+x}{2}} \\
&= \left(\frac{1-\mu}{2}\right)^{\frac{1-(-1)}{2}} \left(\frac{1+\mu}{2}\right)^{\frac{1+(-1)}{2}} + \left(\frac{1-\mu}{2}\right)^{\frac{1-1}{2}} \left(\frac{1+\mu}{2}\right)^{\frac{1+1}{2}} \\
&= \frac{1-\mu}{2}\cdot 1 + 1\cdot\frac{1+\mu}{2} \\
&= \frac{1-\mu+1+\mu}{2} \\
&= 1
\end{aligned}
$$

Mean

$$
\begin{aligned}
\mathbb{E}[X] &= \sum_{\{-1,1\}} x\left(\frac{1-\mu}{2}\right)^{\frac{1-x}{2}} \left(\frac{1+\mu}{2}\right)^{\frac{1+x}{2}} \\
&= -1\cdot\frac{1-\mu}{2} + 1\cdot\frac{1+\mu}{2} \\
&= \frac{-1+\mu+1+\mu}{2} \\
&= \mu
\end{aligned}
$$

Variance

$$
\begin{aligned}
\mathbb{E}\left[X^2\right] &= \sum_{\{-1,1\}} x^2\left(\frac{1-\mu}{2}\right)^{\frac{1-x}{2}} \left(\frac{1+\mu}{2}\right)^{\frac{1+x}{2}} \\
&= 1\cdot\frac{1-\mu}{2} + 1\cdot\frac{1+\mu}{2} \\
&= 1
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{V}[X] &= \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 \\
&= 1 - \mu^2
\end{aligned}
$$

Entropy

$$
\begin{aligned}
\mathrm{H}[X] &= \mathbb{E}\left[\log \frac{1}{p_{X;\mu}}\right] \\
&= \sum_{x \in \{-1,1\}} \left(\log \frac{1}{p_{X;\mu}(x)}\right) p_{X;\mu}(x) \\
&= \left(\log \frac{2}{1-\mu}\right) \frac{1-\mu}{2} + \left(\log \frac{2}{1+\mu}\right) \frac{1+\mu}{2} \\
&= (\log(2) - \log(1-\mu))\frac{1-\mu}{2} + (\log(2) - \log(1+\mu))\frac{1+\mu}{2} \\
&= \log(2)\frac{1-\mu+1+\mu}{2} - \log(1-\mu)\frac{1-\mu}{2} - \log(1+\mu)\frac{1+\mu}{2} \\
&= \log(2) - \log(1-\mu)\frac{1-\mu}{2} - \log(1+\mu)\frac{1+\mu}{2}
\end{aligned}
$$

3.3

$$
\begin{aligned}
\binom{N}{m} + \binom{N}{m-1} &= \frac{N!}{(N-m)!m!} + \frac{N!}{(N-(m-1))!(m-1)!} \\
&= \frac{N!(N-(m-1))!(m-1)! + N!(N-m)!m!}{(N-m)!m!(N-(m-1))!(m-1)!} \\
&= \frac{N!\left((N-m+1)!(m-1)! + (N-m)!m!\right)}{(N-m)!m!(N+1-m))!(m-1)!} \\
&= \frac{N!\cancel{(N-m)!}\cancel{(m-1)!}\left((N-m+1)+m\right)}{\cancel{(N-m)!}m!(N+1-m))!\cancel{(m-1)!}} \\
&= \frac{N!(N+1)}{(N+1-m)!m!} \\
&= \frac{(N+1)!}{((N+1)-m)!m!} \\
&= \binom{N+1}{m}
\end{aligned}
$$

$\mathbf{N = 0}$

$$\begin{aligned}
\sum_{m=0}^{0} \binom{0}{m} x^m &= \binom{0}{0} x^0 \\
&= \frac{0!}{(0-0)!0!} \cdot 1 \\
&= 1
\end{aligned}$$

$\mathbf{N \to N+1}$

$$\begin{aligned}
(1+x)^{N+1} &= (1+x)(1+x)^N \\
&= (1+x) \sum_{m=0}^{N} \binom{N}{m} x^m \qquad \text{induction hypothesis} \\
&= \sum_{m=0}^{N} \binom{N}{m} x^m + \sum_{m=0}^{N} \binom{N}{m} x^{m+1} \\
&= \binom{N}{0} x^0 + \sum_{m=1}^{N} \binom{N}{m} x^m \\
&\quad + \sum_{m=0}^{N-1} \binom{N}{m} x^{m+1} + \binom{N}{N} x^{N+1} \\
&= 1 \cdot x^0 + \sum_{m=1}^{N} \binom{N}{m} x^m \\
&\quad + \sum_{m=1}^{N} \binom{N}{m-1} x^m + 1 \cdot x^{N+1} \\
&= 1 \cdot x^0 + \sum_{m=1}^{N} \left(\binom{N}{m} + \binom{N}{m-1} \right) x^m + 1 \cdot x^{N+1} \\
&= \binom{N+1}{0} x^0 + \sum_{m=1}^{N} \binom{N+1}{m} x^m \\
&\quad + \binom{N+1}{N+1} x^{N+1} \\
&= \sum_{m=0}^{N+1} \binom{N+1}{m} x^m
\end{aligned}$$

Normalization

$$\begin{aligned}
\sum_{m=0}^{N} \binom{N}{m} \mu^m (1-\mu)^{N-m} &= (1-\mu)^N \sum_{m=0}^{N} \binom{N}{m} \left(\mu(1-\mu)^{-1}\right)^m \\
&= (1-\mu)^N \left(1+\mu\,(1-\mu)^{-1}\right)^N \qquad \text{binom. thm.} \\
&= \left((1-\mu)\left(1+\mu\,(1-\mu)^{-1}\right)\right)^N \\
&= \left((1-\mu)+(1-\mu)\mu\,(1-\mu)^{-1}\right)^N \\
&= (1-\mu+\mu)^N \\
&= 1
\end{aligned}$$