Making use of the constraint (3.153), the multinomial distribution in this representation then becomes

$$\begin{aligned}
&\exp\left\{\sum_{k=1}^{M} x_k \ln \mu_k\right\} \\
&= \exp\left\{\sum_{k=1}^{M-1} x_k \ln \mu_k + \left(1-\sum_{k=1}^{M-1} x_k\right)\ln\left(1-\sum_{k=1}^{M-1}\mu_k\right)\right\} \\
&= \exp\left\{\sum_{k=1}^{M-1} x_k \ln\left(\frac{\mu_k}{1-\sum_{j=1}^{M-1}\mu_j}\right) + \ln\left(1-\sum_{k=1}^{M-1}\mu_k\right)\right\}. \qquad (3.155)
\end{aligned}$$

We now identify

$$\ln\left(\frac{\mu_k}{1-\sum_j \mu_j}\right) = \eta_k, \qquad (3.156)$$

which we can solve for μ_k by first summing both sides over k and then rearranging and back-substituting to give

$$\mu_k = \frac{\exp(\eta_k)}{1+\sum_j \exp(\eta_j)}. \qquad (3.157)$$

This is called the *softmax* function or the *normalized exponential*. In this representation, the multinomial distribution therefore takes the form

$$p(\mathbf{x}|\boldsymbol{\eta}) = \left(1+\sum_{k=1}^{M-1}\exp(\eta_k)\right)^{-1}\exp(\boldsymbol{\eta}^{\mathrm{T}}\mathbf{x}). \qquad (3.158)$$

This is the standard form of the exponential family, with parameter vector $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_{M-1})^{\mathrm{T}}$ in which

$$\begin{aligned}
\mathbf{u}(\mathbf{x}) &= \mathbf{x} & (3.159)\\
h(\mathbf{x}) &= 1 & (3.160)\\
g(\boldsymbol{\eta}) &= \left(1+\sum_{k=1}^{M-1}\exp(\eta_k)\right)^{-1}. & (3.161)
\end{aligned}$$

Finally, let us consider the Gaussian distribution. For the univariate Gaussian, we have

$$\begin{aligned}
p(x|\mu,\sigma^2) &= \frac{1}{(2\pi\sigma^2)^{1/2}}\exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\} & (3.162)\\
&= \frac{1}{(2\pi\sigma^2)^{1/2}}\exp\left\{-\frac{1}{2\sigma^2}x^2+\frac{\mu}{\sigma^2}x-\frac{1}{2\sigma^2}\mu^2\right\}, & (3.163)
\end{aligned}$$

Exercise 3.35

which, after some simple rearranging, can be cast in the standard exponential family form (3.138) with