which we can solve for μ to give $\mu = \sigma(\eta)$, where

$$\sigma(\eta) = \frac{1}{1 + \exp(-\eta)} \tag{3.143}$$

is called the *logistic sigmoid* function. Thus, we can write the Bernoulli distribution using the standard representation (3.138) in the form

$$p(x|\eta) = \sigma(-\eta)\exp(\eta x) \tag{3.144}$$

where we have used $1 - \sigma(\eta) = \sigma(-\eta)$, which is easily proved from (3.143). Comparison with (3.138) shows that

$$u(x) = x \tag{3.145}$$
$$h(x) = 1 \tag{3.146}$$
$$g(\eta) = \sigma(-\eta). \tag{3.147}$$

Next consider the multinomial distribution which, for a single observation $\mathbf{x}$, takes the form

$$p(\mathbf{x}|\boldsymbol{\mu}) = \prod_{k=1}^{M} \mu_k^{x_k} = \exp\left\{\sum_{k=1}^{M} x_k \ln \mu_k\right\} \tag{3.148}$$

where $\mathbf{x} = (x_1, \ldots, x_M)^{\mathrm{T}}$. Again, we can write this in the standard representation (3.138) so that

$$p(\mathbf{x}|\boldsymbol{\eta}) = \exp(\boldsymbol{\eta}^{\mathrm{T}}\mathbf{x}) \tag{3.149}$$

where $\eta_k = \ln \mu_k$, and we have defined $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_M)^{\mathrm{T}}$. Again, comparing with (3.138) we have

$$\mathbf{u}(\mathbf{x}) = \mathbf{x} \tag{3.150}$$
$$h(\mathbf{x}) = 1 \tag{3.151}$$
$$g(\boldsymbol{\eta}) = 1. \tag{3.152}$$

Note that the parameters η_k are not independent because the parameters μ_k are subject to the constraint

$$\sum_{k=1}^{M} \mu_k = 1 \tag{3.153}$$

so that, given any $M - 1$ of the parameters μ_k, the value of the remaining parameter is fixed. In some circumstances, it will be convenient to remove this constraint by expressing the distribution in terms of only $M - 1$ parameters. This can be achieved by using the relationship (3.153) to eliminate μ_M by expressing it in terms of the remaining $\{\mu_k\}$ where $k = 1, \ldots, M - 1$, thereby leaving $M - 1$ parameters. Note that these remaining parameters are still subject to the constraints

$$0 \leqslant \mu_k \leqslant 1, \qquad \sum_{k=1}^{M-1} \mu_k \leqslant 1. \tag{3.154}$$