

	Linear Regression	Binomial	One of Many
$P(d x, \omega)$	$K \exp\left(\frac{-\sum_j (d_j - y_j)^2}{2\sigma^2}\right)$	$\prod_j y_j^{d_j} (1 - y_j)^{1-d_j}$	$\prod_j y_j^{d_j}$
Interpretation of output $y$ given $x$	mean value of $d$ conditioned on $x$ $\langle d x \rangle$	$y_j =$ probability the $j^{\text{th}}$ output is 1	$y_j =$ probability that input $x$ is in class $j$
Cost Function $E = -\ln P$	$\frac{1}{2\sigma^2} \sum_{i,j} (d_j^{(i)} - y_j^{(i)})^2$	$-\sum_{i,j} (d_j^{(i)} \ln y_j^{(i)} + (1 - d_j^{(i)}) \ln(1 - y_j^{(i)}))$	$-\sum_{i,j} d_j^{(i)} \ln y_j^{(i)}$
Transfer Function $y = F(z)$	linear $\sigma^2 z$	sigmoid $\frac{1}{1+e^{-z}}$	softmax $F_i(z) = \frac{e^{z_i}}{\sum_j e^{z_j}} = y_i$
$-\frac{\partial E}{\partial z} =$	$(d - y)$	$(d - y)$	$(d - y)$
$F'(z)$	$\sigma^2$	$F(z)(1 - F(z)) = y(1 - y)$	$\frac{\partial F_i(z)}{\partial z_k} = F_i(\delta_{ik} - F_k) = y_i(\delta_{ik} - y_k)$

where

$$\begin{aligned}
x^{(i)} &= i^{\text{th}} \text{ input pattern} \\
&= (x_1^{(i)}, x_2^{(i)}, \dots, x_n^{(i)}) \\
y^{(i)} &= \text{network output associated with input } x^{(i)} \\
&= (y_1^{(i)}, y_2^{(i)}, \dots, y_m^{(i)}) \\
\omega_j &= (\omega_{j1}, \dots, \omega_{jn}) \\
z_j &= \omega_j \cdot x + \beta_j \\
\delta_{ij} &= \text{kroncker delta}
\end{aligned}$$