3a) $p(y;\lambda) = \frac{e^{-\lambda}\lambda^{y}}{y!}$

general GLM: $p(y;\eta) = b(y)\exp(\eta^{T} T(y) - a(\eta))$

$\Rightarrow$ Poisson: $\frac{e^{-\lambda}\lambda^{y}}{y!} = \cancel{\frac{\lambda^{y}}{y!} \cdot \exp(-\lambda)}$

$$= (y!)^{-1} \cdot \exp(\log(e^{-\lambda}) + \log(\lambda^{y}))$$

$$= (y!)^{-1} \cdot \exp(-\lambda + y \cdot \log(\lambda))$$

$$= (y!)^{-1} \cdot \exp(y \cdot \log(\lambda) - \lambda)$$

$\Rightarrow \boxed{b(y) = (y!)^{-1}}$ $\cancel{(w/\ \eta = \lambda)}$

$\cancel{a(\eta) = \eta}$

$\cancel{T(y) = y}$

$\boxed{T(y) = y}$

$\boxed{\eta = \log(\lambda)}$

$\Rightarrow e^{\eta} = \lambda$

$\Rightarrow \boxed{a(\eta) = e^{\eta}}$

$$\int b(y) \exp(\eta \cdot y - a(\eta)) = 1$$

$$' = b(y) \cdot \exp(\eta y - a(\eta)) \cdot (y - a'(\eta))$$

$$a(\eta) = \log\left(\int b(y) \exp(\eta \cdot y)\right)$$

$$\cancel{\int b(y) \exp(\eta \cdot y) \cdot \frac{1}{\int b(y) \exp(\eta \cdot y)}}$$

$$\Rightarrow \int b(y) \exp(\eta y - a(n)) = 1$$

$$\Rightarrow \int b(y) \cdot e^{ny} \cdot e^{-a(\eta)}$$

$$\Rightarrow e^{-a(n)} \int b(y) \cdot e^{ny} = 1$$

$$\Rightarrow e^{a(n)} = \int b(y)\, e^{ny}$$

$$\Rightarrow e^{a(\eta)} \cdot a'(\eta) = b(y) \cdot e^{ny}$$

$$\Rightarrow a'(n) = b(y) \cdot e^{ny} \cdot e^{-a(n)}$$

$$e^{a(n)} \cdot a'(n) = \left(\int_y b(y) \cdot e^{ny}\, dy\right) \frac{d}{dn}$$

$$\uparrow$$

$$\not\equiv$$

$$\cancel{\left(\int_y b(y) \cdot e^{y}\, dy\right) \cdot e^{n}}$$

$$\int_y b(y) \cdot e^{ny - a(n)} \cdot y \, dy$$

$$\forall \qquad \downarrow$$

$$\Rightarrow \quad a'(n) = \cancel{\int_y b(y) \cdot e^{ny - a(n)}}$$

$$= E(y)$$

Mistakes made: - Differentiated by wrong var
- Diff'd integral incorrectly

$$\frac{x}{1+x} = \frac{x(1-x)}{1-x^2} = \frac{x-x^2}{1-x^2} = \frac{1+x-x^2-1}{1-x^2}$$

$$x \cdot (1+x)^{-1}$$

$$= \frac{1-x^2}{1-x^2} + \frac{x-1}{1-x^2}$$

$$= 1 + \frac{(x-1)}{(1-x)(1+x)}$$

$$= 1 - \frac{(\cancel{1-x})}{(\cancel{1-x})(1+x)}$$

$$= 1 - \frac{1}{(1+x)}$$

d) $h(x) = E[y|x] \quad \eta = \theta^T x$

$p(y;\eta) = b(y) \exp(\eta^T T(y) - a(\eta))$

$T(y) = y$

$\Rightarrow \log p(y|x,\theta) =$

$= \log(y) + \eta^T T(y) - a(\eta)$

$= \log(y) + (\theta^T x)^T \cdot y - a(\theta^T x)$

$\Rightarrow \ell_{\theta_j}(\theta) = x_j \cdot y - a'(\theta^T x) \cdot x_j$

$\Rightarrow \theta_i = \theta_i - \alpha \left(a'(\theta^T x) - y \right) x_i$

$\Rightarrow a'(\theta^T x) = h(x) \Rightarrow$ why?

$a'(\theta^T x) = E(T(y)) = h(x)$ (Property of log partition function)

and can be shown because: