a) 
$$\times \sim N(\mu_1 b^2)$$

$$f(x) = \frac{1}{\sigma_1 2 a}$$

a.) 
$$\times \sim N(\mu_1 b^2)$$

$$f(x) = \frac{1}{\sigma_1 2\pi_1} e^{-\frac{1}{\sigma_1}}$$

a.) 
$$\times \sim N(\mu_1 \sigma^2)$$
  

$$f(x) = \frac{1}{\sigma \cdot 2\pi} ex$$

a.) 
$$\times \sim N(\mu_1 b^2)$$
  
 $f(x) = \frac{1}{a \cdot 2\pi} e^{-\frac{1}{2} x}$ 



a.) 
$$\frac{1}{x}$$
  $\frac{1}{x}$   $\frac{1}{x}$   $\frac{1}{x}$   $\frac{1}{x}$   $\frac{1}{x}$   $\frac{1}{x}$   $\frac{1}{x}$   $\frac{1}{x}$ 

Exponential Families

a.) 
$$X \sim N(\mu_1 \sigma^2)$$
 $f(x) = \frac{1}{2} \exp \frac{1}{2}$ 

- f(x)= = exp{= (x=)2}
- $= \exp \left\{ \frac{-1}{2\sigma^2} (x \mu)^2 + \log \left( \frac{1}{\sigma \sqrt{2\pi}} \right) \right\} = \exp \left\{ \frac{-1}{2\sigma^2} x^2 + \log \left( \frac{1}{\sigma \sqrt{2\pi}} \right) + \frac{x \mu \frac{1}{2} \mu^2}{\sigma^2} \right\}$  $a(\sigma^2) = \sigma^2$
- $b(\mu) = \frac{1}{2}\mu^{2}$   $c(x_{1}\sigma^{2}) = \frac{1}{2\sigma^{2}}x^{2} + \log\left(\frac{1}{\sigma\sqrt{2\pi}}\right)$ 

  - Y=Z/N where Z~bin(N,P)  $f_{\gamma}(y) = f_{z}(Ny) \left| \frac{1}{N} \right| = \left( \frac{N}{Ny} \right) p^{Ny} \left( 1 - p \right)^{N-Ny} \frac{1}{N}$
  - = (N) 1/ p Ny (1-p) N- Ny = exp{ log (N) - log(N+ Ny log P+Nlog (1-P)-NY log (1-P) = exp{y (Nlog (1-P) - Nlog(1-P) - log(N)+log (Ny)}

 $b(\lambda+k) = \frac{k!}{y_k c_{-y}} = \frac{\exp\{\log(k!)\}}{\exp\{\log(k!)\}} = \exp\{\log(y_{-y})\}$ 

Show that  $E[s(\theta)] = 0$ . Let  $f(x,\theta)$  be the likelihood of the vector of n observations.

 $= \int_{\mathcal{K}} \left( \frac{1}{f(x_j \theta)} \frac{3}{3\theta} f(x_j \theta) \right) f(x_j \theta) dx \qquad \text{by chain rule.}$ 

Show that  $I(\theta) \equiv Var[s(\theta)] = -E[H(\theta)]$  where  $H(\theta)$  is Hassian matrix of log-likelihood

 $=\frac{\partial}{\partial x}\int_{-\infty}^{\infty}f(x)\phi dx=\frac{\partial}{\partial x}(1)=0.$ 

 $0 = \frac{\partial}{\partial \theta^i} \left[ E[s(\theta)] = \frac{\partial}{\partial \theta^i} \int_{\gamma} \frac{\partial}{\partial \theta^i} \log f(\theta_j X) f(x_j \theta) dx \right]$ 

=  $\int_{V} \frac{\partial}{\partial \theta} f(x_j \theta) dx$  then interchange integral and partial

 $= \int_{\mathcal{X}} \frac{1}{9696} \log f(\theta) x \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \log f(\theta) x \int_$ 

0=NIOG(1=0) b(0) = N109 (17)

c(4) \$) = - [09(4!)

 $S(\theta) = \frac{\partial}{\partial \theta} \log L(\theta) = \frac{\partial}{\partial \theta} = \frac{\partial}{\partial \theta} = \frac{\partial}{\partial \theta} \log f(\theta; \theta)$ 

n= E[s(θ)]

 $\mathbb{E}[a(\theta)] = \int \sqrt{\frac{3\theta}{2}} \log^{\frac{1}{2}}(x;\theta) dx$ 

- c(y, p) = 109( N) (09(N)

- - $\theta = \log \lambda$   $a(\phi) = 1$

$$= E\left[\frac{\partial^{2}}{\partial \sigma \partial \theta^{T}} \log f(X)\theta\right] + E\left[\frac{\partial}{\partial \theta} \log f(\theta)X\right] \underbrace{\frac{\partial}{\partial \theta^{T}} \log f(X)\theta}_{S(\theta)}\right]$$

$$= E\left[H(\theta)\right] + E\left[S(\theta)S(\theta)^{T}\right] \qquad \text{where} \quad H(\theta) = 2nd \text{ derivative of log likelihood.}$$

$$\Rightarrow \text{Var}[S(\theta)] = -E[H(\theta)]$$

$$Var[s(\theta)] = -E[H(\theta)]$$

$$\frac{\partial}{\partial u} = \exp\left\{\frac{a(\phi)}{a(\phi)} + c(d)\phi\right\}$$

c) 
$$f(y) \theta_i \phi$$
 =  $exp\left\{\frac{y\theta - b(\theta)}{a(\phi)} + c(y) \theta\right\}$ 

c) 
$$f(y; \theta, \phi) = exp\left(\frac{y\theta - b(\phi)}{a(\phi)} + c(y; \theta)\right)$$
  
 $S(\theta) = \frac{\lambda}{\lambda \theta} \log f(y; \theta, \phi)$ 

$$S(\theta) = \frac{1}{200} \log f(y; \theta, \emptyset)$$

$$= \frac{1}{200} \log \exp\left(\frac{y\theta - b(\theta)}{a(0)} + c(y, \emptyset)\right)$$

$$= \frac{\partial}{\partial \theta} \left[ \log \exp \left( \frac{a(\theta)}{4\theta - b(\theta)} + c(\theta) \right) \right]$$

$$= \frac{\partial}{\partial \theta} \left[ \log \exp \left( \frac{a(\theta)}{4\theta - b(\theta)} + c(\theta) \right) \right]$$

$$= \frac{3\theta}{3\theta} \left( \frac{\alpha(\phi)}{4\theta - \beta(\phi)} + c(\phi) \right)$$

$$= \frac{y - b'(\theta)}{a(\phi)}$$

$$\frac{1}{3} \frac{1}{a(\phi)}$$

$$\frac{1}{3} \frac{1}{a(\phi)} = 0 \text{ then } E$$

$$E\left[\frac{4-b'(\theta)}{a(\phi)}\right] = 0 \quad \text{then} \quad E(Y) = b'(\theta)$$

$$\operatorname{Aon}\left[\frac{\operatorname{O}(A)}{\operatorname{A}^{-}\operatorname{P}_{1}(\Theta)}\right] = -\operatorname{E}\left[\operatorname{H}(\Theta)\right] \qquad \qquad \frac{9\Theta}{9} \frac{\operatorname{O}(\Phi)}{\operatorname{A}^{-}\operatorname{P}_{1}(\Psi)}$$

$$= \frac{9\Theta}{9} \frac{\operatorname{O}(\Phi)}{\operatorname{A}^{-}\operatorname{P}_{1}(\Psi)}$$

$$= \frac{9\Theta}{9} \operatorname{Ind}\left(\operatorname{A}(A)\right)$$

$$\int_{\partial a_{i}(b)} \int_{\partial a_{i}(b)} \int_{\partial$$

$$\frac{\sqrt{|\alpha(\psi)|^2}}{\alpha(\phi)^2} = \frac{b^*(\theta)}{\alpha(\phi)}$$

$$\sqrt{|\alpha(\psi)|^2} = \alpha(\phi)b^*(\theta)$$

$$Au_{\alpha}(\lambda) = \alpha(\theta) p_{\alpha}(\theta) = \rho_{\alpha} \left( \frac{3w_{\alpha}}{2} \frac{1}{3} m_{\alpha} \right)$$

$$Au_{\alpha}(\lambda) = \alpha(\theta) p_{\alpha}(\theta) = \rho_{\alpha} \left( \frac{3w_{\alpha}}{2} \frac{1}{3} m_{\alpha} \right)$$

$$|a_{\alpha}(y)| = a(\beta) |a_{\alpha}(y)| + a(\beta) |a_{\alpha}(y)|$$

Greneralized Linear Models:  

$$\int f(y_i) \theta_i, \emptyset) = \exp\left\{\frac{y_i \theta_i - b(\theta_i)}{\theta_i w_i} + c\theta_i\right\}$$

a.) 
$$f(y_i; \theta_i, \phi) = \exp\left\{\frac{y_i \theta_i - b(\theta_i)}{\phi / w_i} + c(y_i; \phi/w_i)\right\}$$

$$f(y_i;\theta_i,\theta) = \exp\left\{\frac{y_i\theta_i - b(\theta_i)}{\theta_i w_i} + c_i\right\}$$

$$W_i = P_i(\theta) \qquad d(P_i(\theta)) = \chi_i \theta$$

$$Q(P_i(\theta)) = \chi_i \theta$$

$$\mu = p_i(\theta) \qquad d(p_i(\theta)) = \chi_i \beta$$

$$p_i(\theta_i) = q^{-1}(X_i)$$

$$p_i(\theta_i) = q^{-1}(X_i)$$

P.) A(M) = y

 $p(\theta) = y = exb(\theta)$ b= b1= b1 =expA v= exp{iog(M)}=M

$$\theta' = (p_i)_{-1}(\partial_{-1}(x_i|b))$$

$$\theta'(\theta) = \partial_{-1}(x_i|b)$$

$$\theta'(\theta) = \chi_i|b$$

$$h_1(\theta_1) = \lambda_1(\theta)$$

$$g(p_1(\theta_1)) = \chi_1(p_1)$$

$$h_1(\theta_1) = q_{-1}(\chi_1(p_1))$$

 $Var(Y_i) = \alpha(\emptyset)b^u(\emptyset) = \frac{\phi}{w_i}b^u(\emptyset) = \frac{\phi}{w_i}b^u((\emptyset) = \frac{\phi}{w_i}b^u((b')^{-1}(g^{-1}(X_i^{\dagger}\beta))) = \frac{\phi}{w_i}b^u((b')^{-1}(y_{ki}))$ 

b(e)= log (1+ exp+)

 $(b')^{-1} = \frac{\exp \theta}{1 + \exp \theta}$ 

v(M)=M(1-M)

$$\frac{\partial}{\partial \theta^2} \log f(y;\theta)$$

$$\frac{\partial}{\partial \theta} \frac{y - b'(\theta)}{\partial \theta}$$

 $\lim_{M \to \infty} = \underbrace{\phi}_{V(M)} + \lim_{M \to \infty} V = b^{n}(b^{1-1})$ 

Prug Into variance function  $V(A) = \frac{\exp[\log(\frac{\pi}{A})]}{(1 + \exp[\log(\frac{\pi}{A})])^{2}}$   $= \mu(1-\mu)$ 

(1) canonical link: 
$$g(\mu) = (b')^{-1}(\mu)$$

$$b(\mu) = e^{\mu} = b^{1}(\mu)$$

$$(2) b^{1}(\mu) = \frac{\exp(\mu)}{1 + \exp(\mu)}$$

$$(b^{1})^{-1}(\mu) = \log\left(\frac{1-\mu}{1-\mu}\right)$$

$$y + y \exp(\mu) = \exp(\mu)$$

$$y = \exp(\mu)$$

$$y = \exp(\mu)$$

$$y = \exp(\mu)$$

$$y = \exp(\mu)$$

θ=log λ e<sup>θ=</sup>λ

Fitting GILNC

$$\begin{aligned}
\mathbf{E}_{i}^{T} &= \nabla_{\beta} \log_{\beta}(\beta, \phi) = \nabla_{\beta} \log_{\beta}\left[\prod_{i=1}^{n} \exp\left\{\frac{g_{i}(\beta_{i} - b_{i}(\beta_{i})}{\phi / \omega_{i}} + c_{i}(y_{i}; \phi / \omega_{i})\right\}\right] \\
&= \frac{\lambda}{\lambda \beta} \sum_{i=1}^{n} \frac{y_{i}(\beta_{i} - b_{i}(\beta_{i})}{\phi / \omega_{i}} + c_{i}(y_{i}; \phi / \omega_{i}) \\
&= \sum_{i=1}^{n} \left(\frac{y_{i} - b_{i}(\beta_{i})}{\phi / \omega_{i}}\right) \left(\frac{y_{i}}{\beta ((\omega_{i})^{-1}(y_{i}))}\right) \left(\frac{y_{i}}{\beta (g_{i}^{-1}(y_{i}^{-1}(\beta_{i})))}\right) \\
&= \frac{\lambda}{\lambda \beta} \left(\frac{y_{i} - b_{i}(\beta_{i})}{\phi / \omega_{i}}\right) \left(\frac{y_{i}}{\beta ((\omega_{i})^{-1}(y_{i}))}\right) \left(\frac{y_{i}}{\beta (g_{i}^{-1}(y_{i}^{-1}(\beta_{i})))}\right) \\
&= \frac{\lambda}{\lambda \beta} \left(\frac{y_{i} - b_{i}(\beta_{i})}{\phi / \omega_{i}}\right) \left(\frac{y_{i}}{\beta ((\omega_{i})^{-1}(y_{i}))}\right) \left(\frac{y_{i}}{\beta ((\omega_{i})^{-1}(y_{i}^{-1}(\beta_{i})))}\right) \\
&= \frac{\lambda}{\lambda \beta} \left(\frac{y_{i} - b_{i}(\beta_{i})}{\phi / \omega_{i}}\right) \left(\frac{y_{i}}{\beta ((\omega_{i})^{-1}(y_{i}))}\right) \left(\frac{y_{i}}{\beta ((\omega_{i})^{-1}(y_{i}))}\right) \\
&= \frac{\lambda}{\lambda \beta} \left(\frac{y_{i} - y_{i}(\beta_{i})}{\phi / \omega_{i}}\right) \left(\frac{y_{i}}{\beta ((\omega_{i})^{-1}(y_{i}))}\right) \left(\frac{y_{i}}{\beta ((\omega_{i})^{-1}(y_{i}))}\right) \\
&= \frac{\lambda}{\lambda \beta} \left(\frac{y_{i}}{\beta ((\omega_{i})^{-1}(y_{i}))}\right) \left(\frac{y_{i}}{\beta ((\omega_{i})^{-1}(y_{i}))}\right) \left(\frac{y_{i}}{\beta ((\omega_{i})^{-1}(y_{i}))}\right) \\
&= \frac{y_{i}}{\lambda \beta} \left(\frac{y_{i}}{\beta ((\omega_{i})^{-1}(y_{i}))}\right) \left(\frac{y_{i}}{\beta ((\omega_{i})^{-1}(y_{i}))}\right) \left(\frac{y_{i}}{\beta ((\omega_{i})^{-1}(y_{i}))}\right) \\
&= \frac{y_{i}}{\lambda \beta} \left(\frac{y_{i}}{\beta ((\omega_{i})^{-1}(y_{i}))}\right) \left(\frac{y_{i}}{\beta ((\omega_{i})^{-1}(y_{i})}\right) \left(\frac{y_{i}}{\beta ((\omega_{i})^{-1}(y_{i})}\right) \left(\frac{y_{i}}{\beta ((\omega_{i})^{-1}(y_{i})}\right) \left(\frac{y_{i}}{\beta ((\omega_{i})^{-1}(y_{i}))}\right) \left(\frac{y_{i}}{\beta ((\omega_{i})^{-1}(y_{i})}\right) \left(\frac{y_{i}}{\beta ((\omega$$

$$a) \leq (\beta, \phi) \equiv \nabla_{\beta} \log_{\beta}(\beta, \phi) = \nabla_{\beta} \log_{\beta}\left[\prod_{i=1}^{n} \exp\left\{\frac{g_{i}(\beta_{i} - b_{i}(\beta_{i}))}{\phi / \omega_{i}} + c g_{i}; \rho / \omega_{i}\right\}\right]$$

$$= \frac{\lambda}{\lambda \beta} \sum_{i=1}^{n} \frac{g_{i}(\beta_{i} - b_{i}(\beta_{i}))}{\phi / \omega_{i}} + c g_{i}; \rho / \omega_{i}$$

$$\theta_{i} = (b^{i})^{-1} (\mu_{i})$$

$$\frac{\lambda \delta g_{i}}{\lambda \lambda_{i}} = (b^{i})^{-1} (\mu_{i})$$

$$= \sum_{i=1}^{j=1} \left( \frac{m_i (4_i - M_i)}{\infty} \right) \left( \frac{\lambda_i}{3_i(M_i)} \right) \left( \frac{p_n((p_i)_{-1}(M_i))}{p_n((p_i)_{-1}(M_i))} \right) \text{ peranse } p_n(p_i) = M_i \text{ and } M_i = d_{-1}(\chi_i p_i)$$

$$= \sum_{i=1}^{j=1} \left( \frac{m_i (4_i - M_i)}{\infty} \right) \left( \frac{\lambda_i}{3_i(M_i)} \right) \left( \frac{p_n((p_i)_{-1}(M_i))}{p_n((p_i)_{-1}(M_i))} \right) \text{ peranse } p_n(p_i) = M_i \text{ and } M_i = d_{-1}(\chi_i p_i)$$

$$\Lambda = P_n(h_i)_{-1}$$

e" = 4(1+e") N= 64

b.) 
$$g(\mu) = (b')^{-1}(\mu)$$
  
 $g'(\mu) = \frac{1}{b''(b')^{-1}(\mu)} = \frac{1}{V(\mu)}$ 

$$g(\mu) = (b')^{-1}(\mu)$$

$$g'(\mu) = \frac{1}{b''(b')^{-1}(\mu)} = \frac{1}{V(\mu)}$$

$$subctitute g'(\mu) = \frac{1}{V(\mu)}$$

S(
$$\beta, \phi$$
) =  $\sum_{i=1}^{N-1} \frac{w_i(y_i - A_i) Y_i}{\phi v(A_i) g(A_i)}$  substitute  $g(\zeta)$   
=  $\sum_{i=1}^{N-1} \frac{w_i(y_i - A_i) Y_i}{\phi v(A_i) v(A_i)}$ 

$$S(\beta,\phi) = \sum_{i=1}^{n-1} \frac{w_i(y_i - A_i) X_i}{\phi \vee (A_i) \varphi(A_i)} \qquad \text{substitute} \qquad g^1(y_i - A_i) X_i$$

$$= \sum_{i=1}^{n-1} \frac{w_i(y_i - A_i) X_i}{\phi \vee (A_i) \sqrt{A_i}}$$

$$S(\beta, \phi) = \sum_{i=1}^{n} \frac{w_i(y_i - A_i)y_i}{\phi_i(y_i - A_i)y_i}$$

$$= \sum_{i=1}^{n} \frac{w_i(y_i - A_i)y_i}{\phi_i(y_i - A_i)y_i}$$

$$= \sum_{i=1}^{n} \frac{w_i(y_i - x_i) x_i}{\phi}$$

$$= \sum_{i=1}^{n} \frac{w_i(y_i) v_{in}}{\phi}$$

$$= \sum_{i=1}^{n} \frac{w_i(y_i - w_i) x_i}{\phi}$$
composited tink:  $q(\mu) = \log \left(\frac{\mu}{1 - \mu}\right)$ 

c.) canonical link: 
$$g(\mu) = \log \left(\frac{\mu}{1-\mu}\right)$$

connonical link: 
$$g(\mu) = \log \left(\frac{\mu}{1-\mu}\right)$$

$$\phi = \omega; = 1$$

$$\mu := g^{-1}(x^i; \beta)$$

$$g^{-1}(\theta) = \frac{e^{\theta}}{1+e^{\theta}}$$

$$\psi = \frac{e^{y}(1-\mu)}{1-\mu}$$

$$\phi = w_i = 1$$

$$\mu_i = g^{-1}(x_i^{-1}\beta)$$

 $S(\beta,\phi) = \sum_{i=1}^{n} \frac{w_i(y_i - \mu_i) \chi_i}{\phi}$ 

= \frac{1}{2} y;(x;\b) - b(x;\b)

connonical link:  

$$\phi = w_i = 1$$

$$u_i = g^{-1}(x_i^{-1}\beta)$$



d) 
$$S(\beta, \phi) = \frac{1}{2} \frac{w_i(y_i - y_i) Y_i}{\phi}$$
  
 $H(\beta, \phi) = \frac{\lambda}{\partial \beta} S(\beta, \phi)$ 

109 L(Bo)



- $=\frac{\partial}{\partial \beta}\sum_{i=1}^{n-1}\frac{w_{i}(y_{i}-\mu_{i})x_{i}}{\phi}\qquad \frac{\partial}{\partial \mu}\cdot\frac{\partial \mu}{\partial \rho}.$  $=\sum_{i=1}^{n}\left(\frac{-w_i}{\phi}\frac{\chi_i}{\phi}\right)\frac{g^i(q^i(\chi_i^i\beta))}{\chi_i^{-1}}\qquad \mu_i=g^{-1}(\chi_i^{-1}\beta)$
- $= -\sum_{i=1}^{N} \frac{-w_i x_i x_i^i}{\phi g^i(\mu_i)} = -X^i W X \quad \text{where} \quad W = \text{diag} \left( \frac{w_i}{\phi g(\mu_i)} \right)$
- e)  $f(\beta) = f(\beta_0) + \frac{\partial f(\beta_0)}{\partial \beta} (\beta \beta_0) + \frac{1}{2} (\beta \beta_0)^{1} + (\beta_0) (\beta \beta_0)$ 
  - $\frac{\partial b_1}{\partial t(b_0)} = \sum_{i} \frac{\omega_i}{\omega_i} \frac{\rho_i(x_i b_i)}{\rho_i(x_i b_i)} \cdot \frac{\beta_{i,i}(x_i b_i)}{(\lambda_i \lambda_i) \lambda_i} \times 1$

q1(Mi)

= 2'WX

- = Zo'WX(β-βo)-- (β-βo)'X'WX(β-βo)+c\*
- = 30 WXB-= p1X1WXB+B1X1NXB+C\*

- =  $\hat{y}'WX\beta \frac{1}{2}\beta'X'WX\beta + c^*$  if  $\tilde{y}' = \hat{z}' + \beta o'X'$  and

- =  $-\frac{1}{2}g'W\ddot{y} \frac{1}{2}\beta'\chi'W\chi\beta + \ddot{y}W\chi\beta + c^*$  this only changes the constant

  - Di= X:18
- $X_{n+1} = X_n (H)^{-1} * gradient$

=  $-\frac{1}{2}(\tilde{y}-x\beta)^{\dagger}W(\tilde{y}-x\beta)$  complete square

- drog(g1(ui)) Hessian: exp(0)

  1+exp(0)

  - X: N.D  $\mathbf{p} \cdot \mathbf{p}$ 
    - $\chi' M \chi$ (pn)(nn)(n·p)

f.) Newton's method find roots of f

$$X^{u+1} = X^u - \frac{f_1(X^u)}{f_1(X^u)}$$