05/16 ~~EXAM~~ ALSM 2 EXAM NOTES: GLM Steps.

GLM

1. We have independently collected a set of responses y_i aswell as the values for some explanatory variables stored in vector x_i i.e. we have observations: $\{[y_i, x_i]\}$ $i = 1, \dots, N$

2. Response y_i has a distribution $P_{y|\theta_i}(y_i|\theta_i)$ that is a member of the exponential family, indexed by parameter θ and related to expectation of response $\mathbb{E}[y_i]$

3. Model constructed by linking expectation of response $\mathbb{E}[y_i]$ with the linear predictor $x_i^T\beta$
$$g(\mathbb{E}[y_i]) = x_i^T\beta$$
$$\mathbb{E}[y_i] = g^{-1}(x_i^T\beta)$$

4. link function g is a monotonic differentiable function (ensures inverse g^{-1} exists)

5. Estimate β by $\hat{\beta}$ = argmax likelihood or argmax posterior probability

6. $\mathbb{E}[y] = g^{-1}(x^T\beta)$

BERNOULLI

y is a binary variable ~~P(y|θ)~~

$P_{y|\theta}[y=1|\theta] = \theta$ $\quad P_{y|\theta}(y=0|\theta) = 1-\theta$

$P_{y|\theta} = \theta^y(1-\theta)^{1-y}$ $\quad y \in [0,1]$ $\quad \theta \in [0,1]$ $\quad \mathbb{E}[y] = \theta$