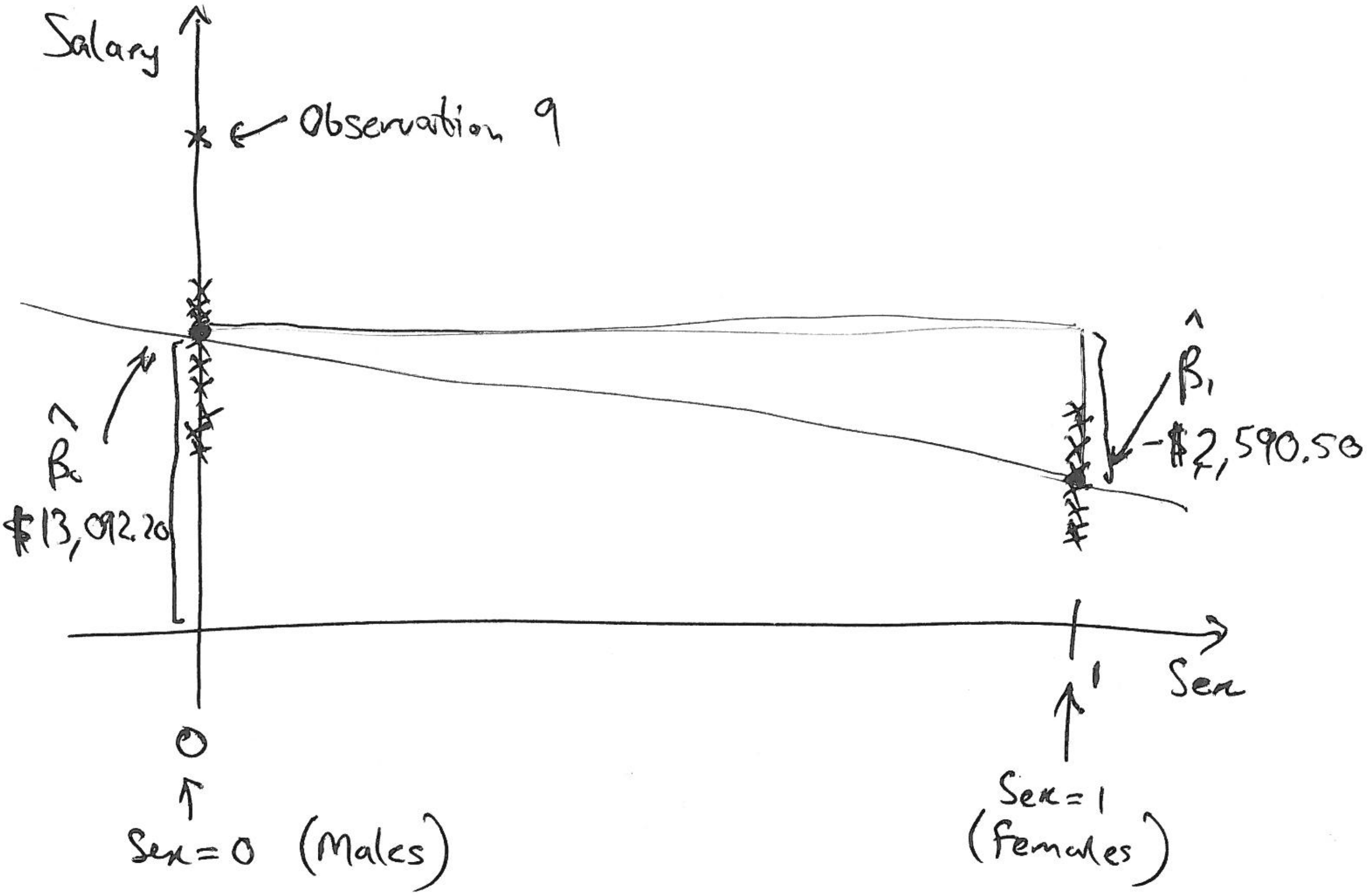
STAT3015/4030/7030 GLMs

1/8/2017

Bank Wages Example



$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X$$
$$= \hat{\beta}_0 + \hat{\beta}_1 D_{sex} \quad \text{where } D_{sex} = \begin{cases} 0 & \text{if Males} \\ 1 & \text{otherwise (females)} \end{cases}$$

for Males $D_{sex} = 0$

$$\hat{Y} = \hat{\beta}_0$$

for Females $D_{sex} = 1$

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 \cdot 1$$
$$= (\hat{\beta}_0 + \hat{\beta}_1)$$

Treatment contrasts / coding

Uses a series of dummy 0/1 indicator variables

$$D_1 = \begin{cases} 1 & \text{if observation in category 1} \\ 0 & \text{otherwise} \end{cases}$$

$$D_2 = \begin{cases} 1 & \text{if obs in cat 2} \\ 0 & \text{otherwise} \end{cases}$$

$\vdots$

D_{k-1} ← we only need $k-1$ dummy variables to code the info for a k level factor variable

The last or k^{th} dummy variable is superfluous – it would be equal to 1 when all the others are 0 !

The category we choose to not have it's own indicator variable is the reference category (in treatment coding this is typically the control group)

Constraint is $D_{\text{ref. cat.}} = 0$

Sum constrasts / coding

Use 0/1/−1 indicator variables

$$I_1 = \begin{cases} 1 & \text{if in category 1} \\ -1 & \text{if in reference category} \\ 0 & \text{otherwise} \end{cases}$$

$$I_2 = \begin{cases} 1 & \text{cat 2} \\ -1 & \text{ref. cat.} \\ 0 & \text{otherwise} \end{cases}$$

$\vdots$

I_{k-1} ← again, we need to choose a ref. category ($= -1 \ \forall \ I_1, \dots I_{k-1}$) & we only need $k-1$ indicator variables

Sum Constraint $\sum\limits_{\text{categories}} n_j \, \text{effect}_j = 0$

suitably weighted for # units in category j → if balanced all weights the same

I's