Baseline Models Difference

As a starting point in Multinomial Logistic Regression we can choose any class to be the **baseline**
Example

- Say we have A, B, C classes
- We pick B to be the baseline
  The model estimate

$$\log\left(\frac{\Pr(\text{Y=A})}{\Pr(Y=B)}\right)=\vec{X}\vec{\beta}_A$$

$$\log\left(\frac{\Pr(\text{Y=C})}{\Pr(Y=B)}\right)=\vec{X}\vec{\beta}_C$$

$\text{After adding exponent for both sides we get}$

$$\frac{P(Y=A)}{P(Y=B)}=e^{\vec{X}\vec{\beta}_A}$$
$$\frac{P(Y=C)}{P(Y=B)}=e^{\vec{X}\vec{\beta}_C}$$

$\text{Let } S_A=e^{\vec{X}\vec{\beta}_A}S_C=e^{\vec{X}\vec{\beta}_C}$

$$P(A)+P(B)+P(C)=1$$

$$P(C)=S_CP(B)$$

$$P(A)=S_AP(B)$$

$$P(B)=P(B)$$

$$S_AP(B)+P(B)+S_CP(B)=1$$

$$P(B)(S_A+S_C+1)=1$$
$$P(B)=\frac{1}{S_A+S_C+1}=\frac{1}{1+e^{\vec{X}\vec{\beta}_A}+e^{\vec{X}\vec{\beta}_C}}$$
$$P(A)=\frac{e^{\vec{X}\vec{\beta}_A}}{1+e^{\vec{X}\vec{\beta}_A}+e^{\vec{X}\vec{\beta}_C}}$$
$$P(C)=\frac{e^{\vec{X}\vec{\beta}_C}}{1+e^{\vec{X}\vec{\beta}_A}+e^{\vec{X}\vec{\beta}_C}}$$

- In the denominator there is $+1$ cause the e
  Why Baseline is needed?
- Cause probabilities must sum to 1

$$P(Y=A)+P(Y=B)+P(Y=C)=1$$

- Only $K-1$ sets of classes are needed to model K model

Generalized

$$\log\left(\frac{P(Y=k|X)}{P(Y=K)|X)}\right)=\vec{X}'\vec{\beta}_k$$

- This is the difference in **log odds** between the class k and the baseline
- The coefficient $\vec{\beta}$ describe the log odds of choosing the class k and over the baseline and how much it will vary from X

$$\log\left(\frac{\Pr(\text{Y=k|X})}{\Pr(Y=K|X)}\right)=\vec{X}\vec{\beta}_k$$

$$\frac{\Pr(Y=k|X)}{\Pr(Y=K|X)}=e^{\vec{X}\vec{\beta}_k}$$

- $\Pr(Y=K|X)=\frac{1}{1\sum e^{\vec{X}\vec{B}}}$ is the baseline class K

- $Pr(Y=k|X)=\frac{e^{\vec{X}\vec{\beta}_k}}{1+\sum e^{\vec{X}\vec{\beta}_k}}$
  In simple terms we subtract **baseline** log odds, calculate the difference between the baseline coefficients and the model coefficients and we get the *relative* coefficients for that model and then plug them into this

$$P(Y=k|X)=\frac{e^{\vec{X}\vec{\beta}_k}}{1+\sum^{K-1} e^{\vec{X}\vec{\beta}}}$$