Part 1:

(a) $p(x_i = j \mid \pi, r) = \binom{x_i + r - 1}{x_i} \pi^{x_i} (1-\pi)^{r}$

$$\mathcal{L}\left[p(x_i \mid \pi, r)\right] = \prod_{i=1}^{N} \binom{x_i + r - 1}{x_i} \pi^{x_i} (1-\pi)^{r}$$

$$\left[\prod_{i=1}^{N} \binom{x_i + r - 1}{x_i}\right] \pi^{\sum x_i} (1-\pi)^{nr}$$

(b) $\hat{\pi}_{ML} = \arg\max_{\pi} \mathcal{L}\left[p(x_i \mid \pi, r)\right]$

$$\Rightarrow \hat{\pi}_{ML} = \arg\max_{\pi} \left[\prod_{i=1}^{N} \binom{x_i + r - 1}{x_i}\right] \pi^{\sum x_i} (1-\pi)^{nr}$$

$$= \nabla_{\pi} \left[\prod_{i=1}^{N} \binom{x_i + r - 1}{x_i}\right] \pi^{\sum x_i} (1-\pi)^{nr} = 0$$

Log-Liklihood

$$\Rightarrow \nabla_{\pi} \prod_{i=1}^{N} \log\left[\binom{x_i + r - 1}{x_i}\right] \pi^{\sum x_i} (1-\pi)^{nr} = 0$$

$$\Rightarrow \nabla_{\pi} \sum_{i=1}^{N} \left[\log\binom{x_i + r - 1}{x_i}\right] + \log \pi^{\sum x_i} + \log(1-\pi)^{nr} = 0$$

2

$$\Rightarrow \nabla_{\pi} \sum_{i=1}^{N}\left[\log\binom{x_i+r-1}{x_i}\right] + \Sigma x_i \log \pi + nr \log(1-\pi) = 0$$

$$\Rightarrow \left[\sum_{i=1}^{N} 0\right] + \frac{\Sigma x_i}{\pi} + \frac{(-nr)}{(1-\pi)} = 0$$

$$\Rightarrow \frac{\Sigma x_i (1-\pi) - \pi nr}{\pi(1-\pi)} = 0$$

$$= \frac{\Sigma x_i - \pi \Sigma x_i - \pi nr}{\pi(1-\pi)} = 0 \Rightarrow \frac{\Sigma x_i - \pi(\Sigma x_i + nr)}{\pi(1-\pi)} = 0$$

$$= \Sigma x_i - \pi(\Sigma x_i + nr) = 0 \quad \Rightarrow \boxed{\hat{\pi}_{ML} = \frac{\Sigma x_i}{\Sigma x_i + nr}}$$

(c) $\hat{\pi}_{MAP} = \operatorname{argmax}_{\pi} \; p(\pi \mid \{x_1 \dots x_n\})$

$$= \operatorname{argmax}_{\pi} \; p \frac{(x_i \mid \pi, r) \cdot p(\pi)}{p(x_i)}$$

$$= \operatorname{argmax}_{\pi} \; \log p(x_i \mid \pi, r) + \log p(\pi) - \log p(x_i)$$

$$= \operatorname{argmax}_{\pi} \prod_{i=1}^{N} \log\left(\binom{x_i+r-1}{x_i} \pi^{x_i} (1-\pi)^{r}\right) + \log\left[\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \pi^{a-1} (1-\pi)^{b-1}\right]$$

$$= \operatorname{argmax}_{\pi} \sum_{i=1}^{N} \log\binom{x_i+r-1}{x_i} + \Sigma x_i \log \pi + nr \log(1-\pi) + \log \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}$$

$$+ (a-1)\log \pi + (b-1)\log(1-\pi)$$

Taking the derivative and equating to zero to maximize π

$$= 0 + \frac{\sum x_i}{\pi} - \frac{nr}{(1-\pi)} + \frac{(a-1)}{\pi} - \frac{(b-1)}{(1-\pi)} = 0$$

$$\frac{\sum x_i + a - 1}{\pi} - \frac{(b-1) + nr}{(1-\pi)} = 0$$

$$\cancel{\sum x_i} + (\sum x_i + a - 1)(1-\pi) - \pi(b - 1 + nr) = 0$$

$$(\sum x_i + a - 1) - (\sum x_i + a - 1)\pi - \pi(b - 1 + nr) = 0$$

$$\sum x_i + a - 1 = \pi\left[(\sum x_i + a - 1) + (b - 1 + nr)\right]$$

$$\boxed{\hat{\pi}_{MAP} = \frac{\sum^{N} x_i + a - 1}{\left(\sum_{i=1}^{n} x_i + a - 1\right) + (nr + b - 1)}}$$

(d) Posterior distribution of Π using Bayes Rule

$$\frac{p(x_i \mid \Pi, r) \cdot p(\Pi)}{\int p(x_i \mid \Pi, r) \cdot p(\Pi)}$$

$$\binom{x_i + r - 1}{x_i} \Pi^{x_i} (1-\Pi)^{r} \quad \frac{\Gamma(a+b)}{\Gamma(a)\,\Gamma(b)} \Pi^{a-1} (1-\Pi)^{b-1}$$

$$\binom{x_i + r - 1}{x_i} \frac{\Gamma(a+b)}{\Gamma(a)\,\Gamma(b)} \Pi^{x_i + a - 1} (1-\Pi)^{r+b-1}$$

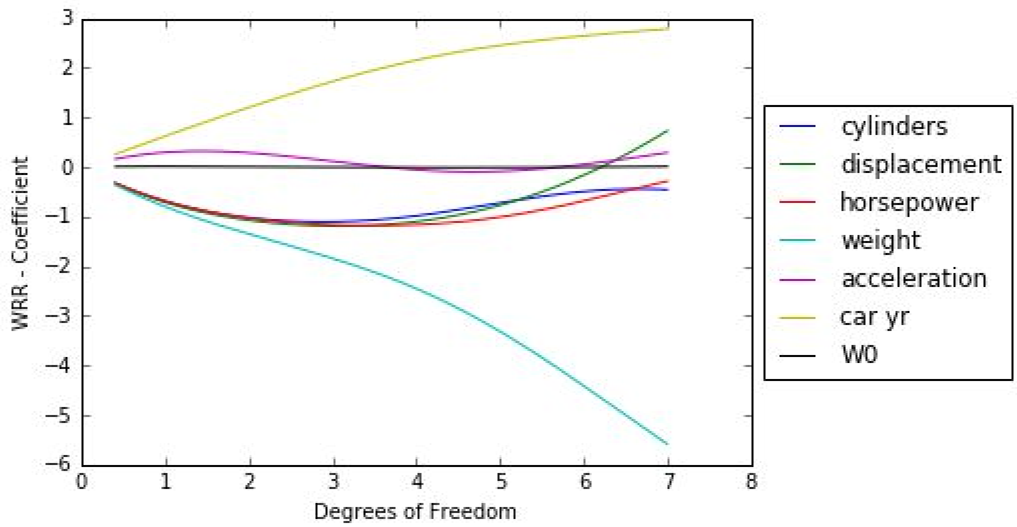
This is a <u>Beta</u> Distribution

(e) $\text{Mean} = \dfrac{x_i + a}{x_i + a + r + b}$ $\quad$ $\text{Variance} = \dfrac{(x_i + a)(r + b)}{(x_i + a)^2 (x + a + r + b + 1)}$

$\hat{\Pi}_{ML} = \hat{\Pi}_{MAP}$ when $a = b = 1$ $\Rightarrow$ The Prior distribution is a <u>Uniform</u> Distribution

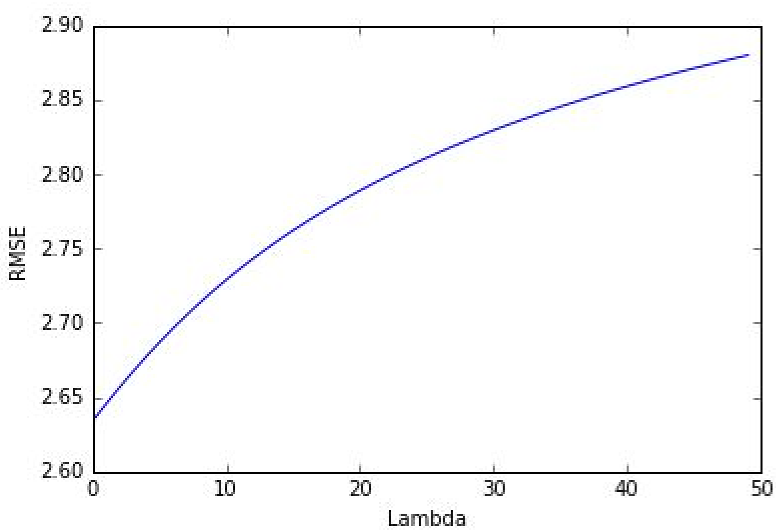
Part -2

a)



b) The 4th Dimension and the 6th Dimension stand out that means that per unit change in **Weight** and per unit change in **car year** have the most impact on the output value of y . Also these two elements have the largest contribution in the variance of the regression.

c) When choosing λ , the Lower λ seems to give a better performance in terms of the RMSE.



d) Based on the plot below , P2 seem to have the best performance when λ = 23. The ideal value of λ has **increased** compared to what we observed part-c with Order -1 regression.