$$Var(z) = E[z^2] - (E[z])^2$$

$$Var(a+z) = Var(z)$$

$$Var(Y \pm z) = Var(Y) \pm Var(z)$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2} = \frac{S_{xy}}{S_x^2}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$t = \frac{\hat{\beta}_1 - 0}{SE(\hat{\beta}_1)}$$

$$t_{n-p-1, \alpha/2}$$

$$RSS = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

$$RSE = \sqrt{\frac{1}{n-2} RSS}$$

(Residual Standard Error)

$$SE(\hat{\beta}_1)^2 = \frac{\sigma^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

$$SE(\hat{\beta}_0)^2 = \sigma^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \right]$$

$$TSS = \sum (y_i - \bar{y})^2$$

$$Reg\,SS = \sum (\hat{y}_i - \bar{y})^2$$

$$[\hat{\beta}_1 - 2 \cdot SE(\hat{\beta}_1), \hat{\beta}_1 + 2 \cdot SE(\hat{\beta}_1)]$$

$$TSS = Reg\,SS + RSS$$

$$[\hat{\beta}_1 - t_{n-2, \alpha/2} \cdot SE(\hat{\beta}_1), \hat{\beta}_1 + t_{n-2, \alpha/2} \cdot SE(\hat{\beta}_1)]$$

$$R^2 = \frac{TSS - RSS}{TSS} = 1 - \frac{RSS}{TSS}$$

$1 - \alpha$ = level of confidence

α = level of doubt/significance

$$r = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}} = \frac{S_{xy}}{S_x S_y}$$

F-statistics (Multiple Linear Reg.) $= \frac{(TSS - RSS)/P}{RSS/(n-p-1)} \sim F_{p, n-p-1}$

$$p(X) = \frac{e^{\beta_0 + \beta_1 X}}{1 + e^{\beta_0 + \beta_1 X}}$$

$$\log\left(\frac{p(X)}{1 - p(X)}\right) = \beta_0 + \beta_1 X$$

case control sampling

$$\hat{\beta}_0^* = \hat{\beta}_0 + \log\frac{\pi}{1-\pi} - \log\frac{\tilde{\pi}}{1-\tilde{\pi}}$$

original fraction of positive class

case control sampling fraction

$$P_{Y|X}(y|x) = \Pr(Y=y \mid X=x) = [p(x)]^y [1-p(x)]^{1-y}$$

$$\ell(\beta_0, \beta_1) = \prod_{i=1}^{N} [p(x_i)]^{y_i} [1-p(x_i)]^{1-y_i}$$

$$= \prod_{i=1}^{N} \left(\frac{e^{\beta_0 + \beta_1 x_i}}{1 + e^{\beta_0 + \beta_1 x_i}}\right)^{y_i} \left(\frac{1}{1 + e^{\beta_0 + \beta_1 x_i}}\right)^{1-y_i}$$

$$\text{cross entropy} = \sum_{i=1}^{N} -\{y_i \log p(x_i) + (1-y_i)\log(1-p(x_i))\}$$

$$z = \frac{\hat{\beta}_i - \beta}{SE(\hat{\beta}_i)}$$ Rejection region: $z > z_{\alpha/2}$ or $z < -z_{\alpha/2}$

Logistic regression with multiple classes -

$$\Pr(Y=k \mid X) = \frac{e^{\beta_{0k} + \beta_{1k}X_1 + \cdots + \beta_{pk}X_p}}{\sum_{l=1}^{K} e^{\beta_{0l} + \beta_{1l}X_1 + \cdots + \beta_{pl}X_p}}$$

Bayes theorem -

$$\Pr(Y=k \mid X=x) = \frac{\Pr(X=x \mid Y=k) \cdot \Pr(Y=k)}{\Pr(X=x)}$$

$$= \frac{\pi_k f_k(x)}{\sum_{l=1}^{K} \pi_l f_l(x)}$$

LDA Gaussian density

$$f_K(x) = \frac{1}{\sqrt{2\pi}\sigma_K} e^{-1/2\left(\frac{x-\mu_K}{\sigma_K}\right)^2}$$

$$p_K(x) = \frac{\pi_K f_K(x)}{\sum_{l=1}^{K} \pi_l f_l(x)} \quad \cdots p=1$$

$$\delta_K(x) = \log \pi_K + \frac{\mu_K x}{2\sigma} - \frac{\mu_K^2}{2\sigma^2}$$

$$\text{Sensitivity/Recall/TPR/Hit Rate} = \frac{TP}{P} = \frac{TP}{TP+FN}$$

$$\text{Specificity/TNR} = \frac{TN}{N} = \frac{TN}{TN+FP}$$

$$\text{Precision} = PPV = \frac{TP}{TP+FP}$$

$$NPV = \frac{TN}{TN+FN}$$

$$PRE_{macro} = \frac{PRE_1 + \cdots + PRE_K}{K}$$

$$PRE_{micro} = \frac{TP_1 + \cdots + TP_K}{TP_1 + \cdots TP_K + FP_1 + \cdots + FP_K}$$

$$F = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

$$F_\beta = \frac{\beta^2 + 1}{\frac{\beta^2}{\text{recall}} + \frac{1}{\text{precision}}}$$

Bootstrap -

$$\alpha = \frac{\sigma_Y^2 - \sigma_{XY}}{\sigma_X^2 + \sigma_Y^2 - 2\sigma_{XY}}$$

$$SE_B(\hat{\alpha}) = \sqrt{\frac{1}{B-1} \sum_{r=1}^{B} (\hat{\alpha}^{*r} - \bar{\hat{\alpha}}^{*})^2}$$

Mallow Cp : $Cp = \frac{1}{n}(RSS + 2d\hat{\sigma}^2)$

$$AIC = -2\log L + 2d$$

$$BIC = \frac{1}{n}(RSS + \log(n) d\hat{\sigma}^2)$$

$$\text{Adjusted } R^2 = 1 - \frac{RSS/(n-d-1)}{TSS/(n-1)}$$

Ridge regression (L_2)

Minimize $RSS + \lambda \sum_{j=1}^{p} \beta_j^2$

$$= \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij}\right)^2 + \lambda \sum_{j=1}^{p} \beta_j^2 \rightarrow \|\beta\|_2^2$$

Lasso (L_1)

minimize $RSS + \lambda \|\beta\|_1$

Elastic Net penalty: $\lambda\left[\frac{1}{2}(1-\alpha)\|\beta\|_2^2 + \alpha \|\beta\|_1\right]$