Some Useful Results and Formulae

These results may be used, without proof, in your solutions.

- $S_{xx} = \sum(x_i - \overline{x})^2 = \sum x_i^2 - n\overline{x}^2. \qquad S_{xy} = \sum(x_i - \overline{x})(y_i - \overline{y}) = \sum x_i y_i - n\overline{x}\,\overline{y}.$
- The population and sample correlation coefficients are
$$\rho = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}} \qquad r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$$
- The least squares estimates for the simple linear model $Y = \beta_0 + \beta_1 X + \varepsilon$ are
$$\hat{\beta}_0 = \overline{y} - \hat{\beta}_1\overline{x} \qquad \hat{\beta}_1 = \frac{S_{xy}}{S_{xx}}$$
- An unbiased estimator of the error variance σ^2 in the simple linear model is
$$\hat{\sigma}^2 = \frac{\mathrm{SSE}}{n-2} = \frac{\sum(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2}{n-2}$$
- The standard error of the estimators in the simple linear model are
$$\mathrm{se}\left(\hat{\beta}_0\right) = \hat{\sigma}\sqrt{\frac{1}{n} + \frac{\overline{x}^2}{S_{xx}}} \qquad \mathrm{se}\left(\hat{\beta}_1\right) = \frac{\hat{\sigma}}{\sqrt{S_{xx}}}$$
- For a given value $X = x_0$, the estimator of $\mathrm{E}(Y|X = x_0)$ and its standard error are
$$\hat{\mu}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_0 \qquad \mathrm{se}(\hat{\mu}_0) = \hat{\sigma}\sqrt{\frac{1}{n} + \frac{(x_0 - \overline{x})^2}{S_{xx}}}.$$
- For a given value $X = x_0$, the best predictor of Y and the standard error of prediction are
$$\hat{y}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_0 \qquad \mathrm{se}(\hat{y}_0) = \hat{\sigma}\sqrt{1 + \frac{1}{n} + \frac{(x_0 - \overline{x})^2}{S_{xx}}}.$$
- For multiple regression with design matrix $\boldsymbol{X}$ the least squares estimators are
$$\hat{\boldsymbol{\beta}} = \left(\boldsymbol{X}^t\boldsymbol{X}\right)^{-1}\boldsymbol{X}^t\boldsymbol{Y}.$$
- The variance-covariance matrix of the estimators in a multiple regression is
$$\mathrm{Var}(\boldsymbol{\beta}) = \left(\boldsymbol{X}^t\boldsymbol{X}\right)^{-1}\sigma^2.$$
- The Projection (Hat) matrix is $\boldsymbol{P} = \boldsymbol{X}\left(\boldsymbol{X}^t\boldsymbol{X}\right)^{-1}\boldsymbol{X}^t$.
- An unbiased estimator of σ^2 in a model with p covariates plus an intercept is $\hat{\sigma}^2 = \mathrm{SSE}/(n-p-1)$.
- For a given value $\boldsymbol{X} = \boldsymbol{x}_0$, the estimator of $\mathrm{E}(Y|\boldsymbol{X} = \boldsymbol{x}_0)$ and its standard error are
$$\hat{\mu}_0 = \boldsymbol{x}_0^t\hat{\boldsymbol{\beta}} \qquad \mathrm{se}(\hat{\mu}_0) = \hat{\sigma}\sqrt{\boldsymbol{x}_0^t(\boldsymbol{X}^t\boldsymbol{X})^{-1}\boldsymbol{x}_0}$$
- For a given value $\boldsymbol{X} = \boldsymbol{x}_0$, the best predictor of Y and the standard error of prediction are
$$\hat{y}_0 = \boldsymbol{x}_0^t\hat{\boldsymbol{\beta}} \qquad \mathrm{se}(\hat{y}_0) = \hat{\sigma}\sqrt{1 + \boldsymbol{x}_0^t(\boldsymbol{X}^t\boldsymbol{X})^{-1}\boldsymbol{x}_0}$$
- For any coefficient in a linear model with $p \geqslant 1$ covariates
$$\frac{\hat{\beta}_j - \beta_j}{\mathrm{se}(\hat{\beta}_j)} \sim t_{n-p-1} \qquad j = 0, 1, \ldots p$$

Some Useful Results and Formulae (continued)

- $R^2 \; = \; 1 - \dfrac{\text{SSE}}{\text{SST}}.$
- $R^2_{adj} \; = \; 1 - \dfrac{\text{MSE}}{s_y^2}.$
- To test if a reduced model is sufficient compared to a full model we use the test statistic

$$F \; = \; \frac{(\text{SSE}_{\text{red}} - \text{SSE}_{\text{full}})/(df_{\text{red}} - df_{\text{full}})}{\text{SSE}_{\text{full}}/df_{\text{full}}}$$

  where the df_{full} and df_{red} are the error degrees of freedom in full and reduced models respectively. This statistic has an F distribution with $df_{\text{red}} - df_{\text{full}}$ and df_{full} degrees of freedom if the reduced model is sufficient.

- The Delta method specifies that

$$\text{E}\big(g(Y)\big) \approx g\big(E(Y)\big) \qquad \text{Var}\big(g(Y)\big) \approx \big(g'\big(\text{E}(Y)\big)\big)^2 \, \text{Var}(Y)$$

  where g' is the derivative of the function g.

- The weighted least squares criterion is $\sum w_i \big(y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{2i} - \cdots - \beta_p x_{ip}\big)^2.$
- The weighted least squares estimates are

$$\hat{\boldsymbol{\beta}} \; = \; (\boldsymbol{X}^t \boldsymbol{W} \boldsymbol{X})^{-1} \boldsymbol{X}^t \boldsymbol{W} \boldsymbol{Y}.$$

  where $\boldsymbol{W}$ is a diagonal matrix with the weights w_i on the diagonal.

- For a model $\mathcal{M}$ with p covariates and an intercept,

$$\begin{aligned} C_p &= \frac{\text{SSE}(\mathcal{M})}{\hat{\sigma}^2} + 2(p+1) - n \\ \text{AIC}(\mathcal{M}) &= n \log(\text{SSE}(\mathcal{M})/n) + 2(p+1) \\ \text{BIC}(\mathcal{M}) &= n \log(\text{SSE}(\mathcal{M})/n) + (p+1)\log(n). \end{aligned}$$