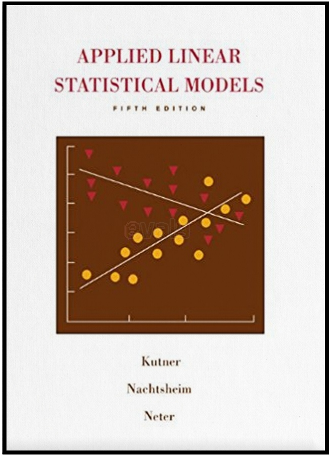
# Applied Linear Statistical Models, Part 1

## **Section 1.8. Normal Error Regression Model**—Proofs of Theorems



# Table of contents

# Theorem 1.8.A

**Theorem 1.8.A.** In the normal error regression model, the likelihood function $L(\beta_0, \beta_1, \sigma^2)$ is minimized at $\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\sigma}^2$ where $\hat{\beta}_0 = b_0$, $\hat{\beta}_1 = b_1$, and $\hat{\sigma}^2 = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n}$ where $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_1$.

**Partial Proof.** Since the likelihood function $L(\beta_0, \beta_1, \sigma^2)$ is positive and the natural logarithm function is increasing, then the maximum of $L = L(\beta_0, \beta_1, \sigma^2)$ and the maximum of $\log(L) = \log(L(\beta_0, \beta_1, \sigma^2))$ occur at the same values of $\beta_0$, $\beta_1$, and $\sigma^2$. We have

$$\log L = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_i)^2.$$

# Theorem 1.8.A

**Theorem 1.8.A.** In the normal error regression model, the likelihood function $L(\beta_0, \beta_1, \sigma^2)$ is minimized at $\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\sigma}^2$ where $\hat{\beta}_0 = b_0$, $\hat{\beta}_1 = b_1$, and $\hat{\sigma}^2 = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n}$ where $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_1$.

**Partial Proof.** Since the likelihood function $L(\beta_0, \beta_1, \sigma^2)$ is positive and the natural logarithm function is increasing, then the maximum of $L = L(\beta_0, \beta_1, \sigma^2)$ and the maximum of $\log(L) = \log(L(\beta_0, \beta_1, \sigma^2))$ occur at the same values of $\beta_0$, $\beta_1$, and $\sigma^2$. We have

$$\log L = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_i)^2.$$

# Theorem 1.8.A (continued 1)

**Partial Proof (continued).** The partial derivatives with respect to the variables $\beta_0$, $\beta_1$, and $\sigma^2$ are

$$\begin{aligned}
\frac{\partial(\log L)}{\partial\beta_0} &= \frac{1}{\sigma^2}\sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_i) \\
\frac{\partial(\log L)}{\partial\beta_1} &= \frac{1}{\sigma^2}\sum_{i=1}^{n}X_i(Y_i - \beta_0 - \beta_1 X_i) \\
\frac{\partial(\log L)}{\partial\sigma^2} &= -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_i)^2.
\end{aligned}$$

# Theorem 1.8.A (continued 2)

**Partial Proof (continued).** Setting each partial equal to zero in order to find critical point(s) we have:

$$\begin{aligned} \sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_i) &= 0 && (*) \\ \sum_{i=1}^{n} X_i(Y_i - \beta_0 - \beta_1 X_i) &= 0 && (**) \\ \frac{\sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_i)}{n} &= \sigma^2 && (***) \end{aligned}$$

Now $(*)$ and $(**)$ are the same as $(*)$ and $(**)$ in the proof of Theorem 1.6.A, and so again we have the solutions to $(*)$ and $(**)$ are $\beta_0 = \hat{\beta}_0$ and $\beta_1 = \hat{\beta}_1$ where...

# Theorem 1.8.A (continued 3)

**Partial Proof (continued).**

$$\hat{\beta}_0 = b_0 = \frac{1}{n}\left(\sum_{i=1}^{n} Y_i - b_1 \sum_{i=1}^{n} X_i\right) \text{ and } \hat{\beta}_1 = b_1 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X})^2}.$$

With $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_i X_i$, we have from $(***)$ that $\hat{\sigma}^2 = \dfrac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{n}$ when $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$.

We still need to confirm that the critical point corresponds a maximum of $L$. To do so, we must address the second derivative test. This requires us to show that the $3^2 \times 3^2 = 9 \times 9$ "Hessian matrix" is negative definite at the critical point. We leave this as an (imposing) exercise. For details on extrema of functions of more than two variables, see my online notes for Vector Analysis (MATH 4317/5317) on Section 3.3. Extrema of Real-Valued Functions (these notes are in preparation). □

# Theorem 1.8.A (continued 3)

**Partial Proof (continued).**

$$\hat{\beta}_0 = b_0 = \frac{1}{n}\left(\sum_{i=1}^{n} Y_i - b_1 \sum_{i=1}^{n} X_i\right) \text{ and } \hat{\beta}_1 = b_1 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X})^2}.$$

With $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_i X_i$, we have from $(***)$ that $\hat{\sigma}^2 = \dfrac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{n}$ when $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$.

We still need to confirm that the critical point corresponds a maximum of $L$. To do so, we must address the second derivative test. This requires us to show that the $3^2 \times 3^2 = 9 \times 9$ "Hessian matrix" is negative definite at the critical point. We leave this as an (imposing) exercise. For details on extrema of functions of more than two variables, see my online notes for Vector Analysis (MATH 4317/5317) on Section 3.3. Extrema of Real-Valued Functions (these notes are in preparation). □