UC Berkeley
Department of Statistics

STAT 210A: Introduction to Mathematical Statistics

Problem Set 10

Fall 2014

Issued: Saturday, November 15 **Due:** Friday, November 21

Problem 10.1
Read Chapter 14 of Keener.

Problem 10.2
Consider the general linear model with normality:

$$Y \sim N(X\beta, \sigma^2 I), \quad \beta \in \mathbb{R}^p, \quad \sigma^2 > 0.$$

If the rank r of X equals p, show that $(\hat{\beta}, S^2)$ is a complete sufficient statistic.

Problem 10.3
Inverse linear regression. Consider the model for simple linear regression:

$$Y_i \sim \beta_1 + \beta_2(x_i - \bar{x}) + \epsilon_i, \qquad i = 1, \ldots, n,$$

studied in Section 14.5.

(a) Derive a level-α test of $H_0 : \beta_2 = 0$ versus $H_1 : \beta_2 \neq 0$.

(b) Let y_0 denote a "target" value for the mean of Y. The regression line $\beta_1 + \beta_2(x - \bar{x})$ achieves this value when the independent variable x equals

$$\theta = \bar{x} + \frac{y_0 - \beta_1}{\beta_0}.$$

Derive a level-α test of $H_0 : \theta = \theta_0$ versus $H_1 : \theta \neq \theta_0$. (Hint: You may want to find a test first assuming $y_0 = 0$. After a suitable transformation, the general case should be similar.)

(c) Use duality to find a confidence region, first discovered by Fieller (1954), for θ. Show that this region is an interval if the test in part (a) rejects $\beta_2 = 0$.

Problem 10.4
Consider a regression version of the two-sample problem in which:

$$Y_i \quad = \quad \begin{cases} \beta_1 + \beta_2 x_i + \epsilon_i & i = 1, \ldots, n_1; \\ \beta_3 + \beta_4 x_i + \epsilon_i & i = n_1 + 1, \ldots, n_1 + n_2 = n, \end{cases}$$

with $\epsilon_i, \ldots, \epsilon_n$ i.i.d. from $N(0, \sigma^2)$. Derive a $1-\alpha$ confidence interval for $\beta_4 - \beta_2$, the difference between the two regression slopes.

Problem 10.5
A variable Y has a log-normal distribution with parameters μ and σ^2 if $\log Y \sim N(\mu, \sigma^2)$.

(a) Find the mean and density for the log-normal distribution.

(b) If $Y_1, \dots, Y_n$ are i.i.d. from the log-normal distribution with unknown parameters μ and σ^2, find the UMVU for μ.

(c) If $Y_1, \dots, Y_n$ are i.i.d. from the log-normal distribution with parameters μ and σ^2, with σ^2 a known constant, find the UMVU for the common mean $\eta = EY_i$.

(d) In simple linear regression, $Y_1, \dots, Y_n$ are independent with $Y_i \sim N(\beta_1 + \beta_2 x_i)$. In some applications this model may be inappropriate because the Y_i are positive; perhaps Y_i is the weight or volume of the ith unit. Suggest a similar model without this defect based on the log-normal distribution. Explain how you would estimate β_1 and β_2 in your model.