From Casella and Berger, do the following problems from Chapter 6:

Homework 2: 8, 9, 13, and 16.
Homework 3: 18, 20, 22, 30, and 31.

These are extra problems that I have given on past exams (in STAT 713 or in related courses). You do not have to turn these in.

6.1. Suppose that $X_1, X_2, ..., X_n$ is an iid sample from

$$f_X(x|\theta) = \frac{1}{2\theta} I(0 < x < 2\theta),$$

where $\theta > 0$.
(a) Show that $T(\mathbf{X}) = X_{(n)}$ is a minimal sufficient statistic for this family.
(b) Is $T(\mathbf{X})$ is complete? Prove or disprove.

6.2. Suppose that $X_1, X_2, ..., X_n$ is an iid sample from a lognormal distribution with pdf

$$f_X(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma x} e^{-(\ln x - \mu)^2/2\sigma^2} I(x > 0).$$

Find a two-dimensional sufficient statistic for $\boldsymbol{\theta} = (\mu, \sigma^2)'$.

6.3. Consider a toxicology study with k groups of animals who are given a drug at distinct dose levels $d_1, d_2, ..., d_k$, respectively (these are fixed by the experimenter; not random). The animals are monitored for a reaction to the drug. In group i, let n_i (fixed) denote the total number of animals dosed, and let Y_i denote the number of animals that respond to the drug. The observations $Y_1, Y_2, ..., Y_k$ are treated as independent random variables, where $Y_i \sim b(n_i, p_i)$; $i = 1, 2, ..., k$, where p_i is the probability that an individual animal responds to dose d_i. A standard assumption in such toxicology studies is that

$$\ln\left(\frac{p_i}{1 - p_i}\right) = \beta_0 + \beta_1 d_i,$$

for $i = 1, 2, ..., k$, where β_0 and β_1 are real parameters (this is merely logistic regression using dose as a predictor). Find a two-dimensional sufficient statistic for $\boldsymbol{\beta} = (\beta_0, \beta_1)'$.

6.4. Recall that if X has a beta(α, β) distribution, then the probability density function (pdf) of X is

$$f_X(x|\alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha - 1}(1 - x)^{\beta - 1} I(0 < x < 1).$$

where $\alpha > 0$ and $\beta > 0$. In this problem, we are going to consider the beta subfamily where $\alpha = \beta = \theta$. Let $X_1, X_2, ..., X_n$ denote an iid sample from a beta(θ, θ) distribution.
(a) Show that

$$T(\mathbf{X}) = \prod_{i=1}^{n} X_i(1 - X_i)$$

is a sufficient statistic for θ. Is $T(\mathbf{X})$ complete?

(b) The two-dimensional statistic

$$\mathbf{T}^*(\mathbf{X}) = \left(\prod_{i=1}^n X_i, \prod_{i=1}^n (1 - X_i)\right)'$$

is also a sufficient statistic for θ. What must be true about the conditional distribution $f_{\mathbf{X}|\mathbf{T}^*}(\mathbf{x}|\mathbf{t})$?
(c) Show that $\mathbf{T}^*(\mathbf{X})$ is not a complete statistic.

6.5. Suppose that T is a complete statistic. Prove that any function of T is also complete.

6.6. Let Y be a random variable that follows a $\mathcal{U}(0, \theta)$ distribution, where $\theta > 1$ is an unknown parameter. Suppose that Y itself is not observed; instead, the "truncated" version of it

$$X = YI(Y \geq 1) + I(Y < 1)$$

is observed.
(a) Show that the pdf of X is

$$f_X(x|\theta) = \theta^{-1} I(1 < x < \theta) + \theta^{-1} I(x = 1).$$

That is, the distribution of X has pdf given by $\theta^{-1} I(1 < x < \theta)$ and a point mass on $\{x = 1\}$.
(b) Let $X_1, X_2, ..., X_n$ be a random sample from the distribution in part (a). Show that the pdf of the maximum order statistic $X_{(n)}$ is given by

$$f_{X_{(n)}}(x|\theta) = n\theta^{-n} x^{n-1} I(1 < x < \theta) + \theta^{-n} I(x = 1).$$

(c) Show that $X_{(n)}$ is a sufficient statistic for θ. Is $X_{(n)}$ complete?

6.7. Suppose that $X_1, X_2, ..., X_n$ is an iid sample from

$$f_X(x|\theta) = \frac{1}{2\theta} I(0 < x < 2\theta),$$

where $\theta > 0$.
(a) Show that $T = T(\mathbf{X}) = X_{(n)}$ is a minimal sufficient statistic for this family.
(b) Find an ancillary statistic $S = S(\mathbf{X})$ for σ^2. Is $\mathbf{T}$ independent of S?

6.8. Suppose that $X_1, X_2, ..., X_n$ is an iid sample from

$$f_X(x|\theta) = \frac{1}{\theta} x^{(1-\theta)/\theta} I(0 < x < 1),$$

where $\theta > 0$.
(a) Show that $T = T(\mathbf{X}) = \prod_{i=1}^n X_i$ is a sufficient statistic.
(b) Is T complete?

6.9. Suppose that $X_1, X_2, ..., X_n$ are iid from

$$f_X(x|\alpha, \beta) = \frac{1}{\beta} e^{-(x-\alpha)/\beta} I(x > \alpha)$$

where $-\infty < \alpha < \infty$ and $\beta > 0$.
(a) Show that

$$\mathbf{T}(\mathbf{X}) = \begin{pmatrix} T_1(\mathbf{X}) \\ T_2(\mathbf{X}) \end{pmatrix} = \begin{pmatrix} X_{(1)} \\ \overline{X} \end{pmatrix}$$

is a sufficient statistic for $\boldsymbol{\theta} = (\alpha, \beta)'$.
(b) Find the distribution of $T_1(\mathbf{X}) = X_{(1)}$.
(c) Carefully argue that $T_1(\mathbf{X})$ and $S(\mathbf{X}) = \sum_{i=1}^n (X_i - X_{(1)})$ are independent statistics.

6.10. Suppose that $X_1, X_2, ..., X_n$ is an iid sample from the probability density function (pdf) given by

$$f_X(x|\theta) = \theta x^{-2} I(x > \theta),$$

where $\theta > 0$.
(a) Show that $\{f_X(x|\theta) : \theta > 0\}$ is a scale family. Identify the standard pdf and the scale parameter.
(b) Show that $T(\mathbf{X}) = X_{(1)}$ is a sufficient statistic for $\{f_X(x|\theta) : \theta > 0\}$.
(c) I have shown (so you don't have to) that $T(\mathbf{X}) = X_{(1)}$ is a complete statistic. Give a statistic that is independent of $T(\mathbf{X})$. Verify your claim.

6.11. Suppose that $X_1, X_2, ..., X_n$ is an iid sample from

$$f_X(x|\theta) = \frac{-\theta^x}{x \log(1-\theta)}$$

where $x = 1, 2, ...,$ and where $\theta \in (0, 1)$.
(a) Argue that $T = T(\mathbf{X}) = X_1 + X_2 + \cdots + X_n$ is a sufficient for θ.
(b) Is T complete?

6.12. Consider data that follow an exponential regression with no intercept:

$$Y_i \overset{ind}{\sim} \exp(\beta x_i),$$

where the scalar parameter $\beta > 0$ is unknown and the x_i's > 0 are fixed and known for $i = 1, 2, ..., n$. That is, $Y_1, Y_2, ..., Y_n$ are independent random variables with pdfs

$$f_{Y_i}(y) = \frac{1}{\beta x_i} \exp\left(-\frac{y}{\beta x_i}\right) I(y > 0).$$

(a) Find a sufficient statistic $T = T(\mathbf{Y})$.
(b) Find the mean and variance of T.

6.13. Suppose that $X_1, X_2, ..., X_n$ is an iid sample from the probability density function (pdf)

$$f_X(x|\theta) = \frac{2x}{\theta} e^{-x^2/\theta} I(x > 0),$$

where $\theta > 0$. Show that $T = \sum_{i=1}^n X_i^2$ is a complete and sufficient statistic.