1. Suppose that $Z \sim \mathcal{N}(\mu, 1)$ and $V \sim \chi^2_k$. If Z and V are independent, then

$$T = \frac{Z}{\sqrt{V/k}}$$

has a noncentral t distribution with k degrees of freedom and noncentrality parameter μ. For $k > 1$, show that

$$E(T) = \frac{\mu\sqrt{k}\Gamma\{(k-1)/2\}}{\sqrt{2}\Gamma(k/2)},$$

where $\Gamma(\cdot)$ is the usual gamma function.

2. It is desired to estimate the weight of each of four objects on a balance. Let β_1, β_2, β_3, and β_4 be the true weights. Suppose that weight measurements are subject to additive error because of imperfection in the scale, so that for any given object with true weight β, its measured weight is a random variable satisfying

$$Y = \beta + \epsilon,$$

where $E(\epsilon) = 0$ and $\text{var}(\epsilon) = \sigma^2 > 0$. Suppose that each object is weighed separately once. In addition, suppose that the first and second objects are weighed together once and that the third and fourth objects are weighed together once. Thus, we have 6 measurements $Y_1, Y_2, ..., Y_6$ which satisfy

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \\ Y_6 \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_1 + \beta_2 \\ \beta_3 + \beta_4 \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \end{pmatrix},$$

where $E(\epsilon_j) = 0$ and $\text{var}(\epsilon_j) = \sigma^2 > 0$, for $j = 1, 2, ..., 6$, and $\text{cov}(\epsilon_j, \epsilon_k) = 0$, for $j \neq k$. Set $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4)'$.

(a) A full-rank linear model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ applies. Find $\widehat{\boldsymbol{\beta}}$, the BLUE of $\boldsymbol{\beta}$.

(b) For the estimators $\widehat{\beta}_i$, $i = 1, 2, 3, 4$, of part (a), calculate $\text{var}(\widehat{\beta}_i)$.

(c) Suppose that you are told that $\beta_1 - \beta_2 = m_1$ and that $\beta_3 - \beta_4 = m_2$, where m_1 and m_2 are known. Find a linear unbiased estimator of $\boldsymbol{\beta}$, say $\widetilde{\boldsymbol{\beta}}$, that incorporates this additional knowledge.

(d) For $i = 1, 2, 3, 4$, compute the efficiency

$$\text{eff}(\widehat{\beta}_i \text{ to } \widetilde{\beta}_i) \equiv \frac{\text{var}(\widehat{\beta}_i)}{\text{var}(\widetilde{\beta}_i)},$$

thereby calculating the relative increase in precision gained by the additional knowledge.

3. Consider the one-way, fixed-effects ANOVA model

$$Y_{ij} = \mu + \alpha_i + \epsilon_{ij},$$

for $i = 1, 2, ..., a$ and $j = 1, 2, ..., n$. Suppose that ϵ_{ij} are iid $\mathcal{N}(0, \sigma^2)$ random variables. The parameter σ^2 is unknown.
(a) Show that the MLE of $\gamma = \sum_{i=1}^{a} c_i \alpha_i$ is $\widehat{\gamma} = \sum_{i=1}^{a} c_i \overline{Y}_{i+}$, where $\overline{Y}_{i+}$ is the ith sample treatment mean.
(b) Derive a $100(1-\alpha)$ percent confidence interval for γ. Explain as many details as possible.

4. Consider the model

$$Y_{ij} = \mu + \tau_i + \gamma_i(x_{ij} - \overline{x}_{i+}) + \epsilon_{ij},$$

for $i = 1, 2$ and $j = 1, 2, ..., n$. Assume that ϵ_{ij} are iid $\mathcal{N}(0, \sigma^2)$ random variables, where $\sigma^2 > 0$. As a frame of reference, suppose that we are studying the effects of two drugs on weight loss (Y) in the situation wherein a covariable x (e.g., initial weight, etc.) is available for each subject. Assume that $\sum_{j=1}^{n}(x_{ij} - \overline{x}_{i+})^2 > 0$ for $i = 1, 2$ and that $x_{1j} \neq x_{2j}$ for at least one j.
(a) Write the model in matrix form $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, identifying all vectors and matrices.
(b) The researcher believes that $\tau_1 = \tau_2$ and that $\gamma_1 = \gamma_2$. If the researcher is correct, sketch a picture of what you would expect the observed data to look like. Use different plotting symbols for the two drugs.
(c) Suppose that we were interested in testing the hypothesis

$$H_0 : \tau_1 = \tau_2,\ \gamma_1 = \gamma_2$$
versus
$$H_1 :\ \text{not } H_0.$$

Write a reduced model, say $\mathbf{Y} = \mathbf{W}\boldsymbol{\gamma} + \boldsymbol{\beta}$, that applies when H_0 is true. Identify all vectors and matrices in your reduced model.
(d) Describe a procedure, in full detail, how you would test the hypothesis in (c). Be very precise, rigorous, and thorough in your explanation. Your description should end with how the test statistic is formed (you can leave it in terms of matrices and vectors), what its distribution is under H_0, and when we reject H_0.

5. Consider the linear model

$$Y_i = \theta_i + \epsilon_i,$$

for $i = 1, 2, 3, 4$, where $\theta_1 + \theta_2 + \theta_3 + \theta_4 = 0$, and the errors ϵ_i are iid $\mathcal{N}(0, \sigma^2)$. Define the statistics

$$U_1 = (Y_1 - Y_3)^2 \qquad \text{and} \qquad U_2 = (Y_1 + Y_2 + Y_3 + Y_4)^2.$$

(a) Describe, with explanation, the joint distribution of U_1 and U_2.
(b) Show that an appropriate F statistic for testing $H_0 : \theta_1 = \theta_3$ is

$$F = \frac{2U_1}{U_2}.$$

6. An engineer is interested in determining the effectiveness of four adhesive systems for bonding insulation to a chamber. The adhesives were applied both with and without a primer. Tests of peel strength (Y) were conducted on two different thicknesses of rubber. Three replications were recorded for each treatment combination. Here are the data:

	A (Thickness)							
	a_1				a_2			
B (Adhesive)	b_1	b_2	b_3	b_4	b_1	b_2	b_3	b_4
	60	57	20	20	52	73	52	77
with primer (c_1)	63	52	20	53	79	56	33	78
	57	55	20	44	76	57	32	70
	59	51	30	49	78	52	38	77
without primer (c_2)	48	44	32	59	72	42	37	76
	51	42	37	55	72	51	35	79

Consider the linear model

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijkl},$$

for $i = 1, 2, ..., a$, $j = 1, 2, ..., b$, $k = 1, 2, ..., c$, and $l = 1, 2, ..., n$, where $\epsilon_{ijk} \sim$ iid $\mathcal{N}(0, \sigma^2)$, and where $a = 2$, $b = 4$, and $c = 3$. Fit all sensible submodels. Which submodels fit as well as the full three-factor model?

7. Consider the one way, fixed effects ANOVA model

$$Y_{ij} = \mu + \alpha_i + \epsilon_{ij},$$

for $i = 1, 2, ..., 5$ and $j = 1, 2, ..., 20$, where ϵ_{ij} are iid $\mathcal{N}(0, \sigma^2)$ random variables. Set

$$\boldsymbol{\beta} = (\mu, \alpha_1, \alpha_2, ..., \alpha_5)'.$$

Consider the solution to the normal equations $\widehat{\mu} = 0$, $\widehat{\alpha}_i = \overline{Y}_{i+}$, for $i = 1, 2, ..., 5$. Suppose that $\overline{Y}_{1+} = 10$, $\overline{Y}_{2+} = 12$, $\overline{Y}_{3+} = 5$, $\overline{Y}_{4+} = 10$, $\overline{Y}_{5+} = 8$, and $\sum_{i=1}^{5}\sum_{j=1}^{20} Y_{ij}^2 = 12460$.
(a) Test $H_0 : \mu + \alpha_2 = 0$ versus $H_1 : \mu + \alpha_2 > 0$ using $\alpha = 0.05$.
(b) Write the hypothesis $H_0 : \alpha_1 = \alpha_2 = \alpha_3$ in the form $H_0 : \mathbf{K}'\boldsymbol{\beta} = \mathbf{m}$, giving explicit expressions for $\mathbf{K}$ and $\mathbf{m}$. Calculate the F statistic, showing all of your work.

8. Consider the general linear model

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon},$$

where $\mathbf{X}$ is $n \times p$ with rank $r \leq p$, $\boldsymbol{\beta}$ is $p \times 1$, and $\boldsymbol{\epsilon} \sim \mathcal{N}_n(\mathbf{0}, \sigma^2\mathbf{I})$. Derive the likelihood ratio test of $H_0 : \sigma^2 = \sigma_0^2$ versus $H_0 : \sigma^2 \neq \sigma_0^2$.

9. Consider the cell means ANOVA model

$$Y_{ij} = \mu_i + \epsilon_{ij}$$

for $i = 1, 2, 3$ and $j = 1, 2, ..., n$, where ϵ_{ij} are iid $\mathcal{N}(0, \sigma^2)$ random variables. The restriction

$$\mu_3 - \mu_2 = \mu_2 - \mu_1$$

is placed on the parameters. Set $\boldsymbol{\beta} = (\mu_1, \mu_2, \mu_3)'$.
(a) Write this as a general linear model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$. What are $\mathbf{X}$ and $\boldsymbol{\beta}$? How do you express the restriction in the form $\mathbf{P}'\boldsymbol{\beta} = \boldsymbol{\delta}$?
(b) Find the restricted least squares estimator, $\widehat{\boldsymbol{\beta}}_H$. Express this estimator in terms of the overall mean $\overline{Y}_{++}$ and the treatment means $\overline{Y}_{i+}$, for $i = 1, 2, 3$.
(c) Let $Q(\boldsymbol{\beta}) = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$ and let $\widehat{\boldsymbol{\beta}}$ denote the unrestricted least squares estimator. How do $Q(\widehat{\boldsymbol{\beta}})$ and $Q(\widehat{\boldsymbol{\beta}}_H)$ compare?