Chapter 11

Analysis of Variance and Regression

11.1 a. The first order Taylor's series approximation is

$$\mathrm{Var}[g(Y)] \approx [g'(\theta)]^2 \cdot \mathrm{Var}\,Y = [g'(\theta)]^2 \cdot v(\theta).$$

b. If we choose $g(y) = g^*(y) = \int_a^y \frac{1}{\sqrt{v(x)}}\,dx$, then

$$\frac{dg^*(\theta)}{d\theta} = \frac{d}{d\theta}\int_a^\theta \frac{1}{\sqrt{v(x)}}\,dx = \frac{1}{\sqrt{v(\theta)}},$$

by the Fundamental Theorem of Calculus. Then, for any θ,

$$\mathrm{Var}[g^*(Y)] \approx \left(\frac{1}{\sqrt{v(\theta)}}\right)^2 v(\theta) = 1.$$

11.2 a. $v(\lambda) = \lambda$, $g^*(y) = \sqrt{y}$, $\frac{dg^*(\lambda)}{d\lambda} = \frac{1}{2\sqrt{\lambda}}$, $\mathrm{Var}\,g^*(Y) \approx \left(\frac{dg^*(\lambda)}{d\lambda}\right)^2 \cdot v(\lambda) = 1/4$, independent of λ.

b. To use the Taylor's series approximation, we need to express everything in terms of $\theta = \mathrm{E}Y = np$. Then $v(\theta) = \theta(1 - \theta/n)$ and

$$\left(\frac{dg^*(\theta)}{d\theta}\right)^2 = \left(\frac{1}{\sqrt{1-\frac{\theta}{n}}} \cdot \frac{1}{2\sqrt{\frac{\theta}{n}}} \cdot \frac{1}{n}\right)^2 = \frac{1}{4n\theta(1-\theta/n)}.$$

Therefore

$$\mathrm{Var}[g^*(Y)] \approx \left(\frac{dg^*(\theta)}{d\theta}\right)^2 v(\theta) = \frac{1}{4n},$$

independent of θ, that is, independent of p.

c. $v(\theta) = K\theta^2$, $\frac{dg^*(\theta)}{d\theta} = \frac{1}{\theta}$ and $\mathrm{Var}[g^*(Y)] \approx \left(\frac{1}{\theta}\right)^2 \cdot K\theta^2 = K$, independent of θ.

11.3 a. $g_\lambda^*(y)$ is clearly continuous with the possible exception of $\lambda = 0$. For that value use l'Hôpital's rule to get

$$\lim_{\lambda\to 0} \frac{y^\lambda - 1}{\lambda} = \lim_{\lambda\to 0} \frac{(\log y)y^\lambda}{1} = \log y.$$

b. From Exercise 11.1, we want to find $v(\lambda)$ that satisfies

$$\frac{y^\lambda - 1}{\lambda} = \int_a^y \frac{1}{\sqrt{v(x)}}\,dx.$$

Taking derivatives

$$\frac{d}{dy}\left(\frac{y^\lambda - 1}{\lambda}\right) = y^{\lambda-1} = \frac{d}{dy}\int_a^y \frac{1}{\sqrt{v(x)}}\,dx = \frac{1}{\sqrt{v(y)}}.$$

Thus $v(y) = y^{-2(\lambda-1)}$. From Exercise 11.1,

$$\mathrm{Var}\left(\frac{y^\lambda - 1}{\lambda}\right) \approx \left(\frac{d}{dy}\frac{\theta^\lambda-1}{\lambda}\right)^2 v(\theta) = \theta^{2(\lambda-1)}\theta^{-2(\lambda-1)} = 1.$$

Note: If $\lambda = 1/2$, $v(\theta) = \theta$, which agrees with Exercise 11.2(a). If $\lambda = 1$ then $v(\theta) = \theta^2$, which agrees with Exercise 11.2(c).

11.5 For the model

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \quad i = 1, \dots, k, \quad j = 1, \dots, n_i,$$

take $k = 2$. The two parameter configurations

$$\begin{aligned}(\mu, \tau_1, \tau_2) &= (10, 5, 2)\\ (\mu, \tau_1, \tau_2) &= (7, 8, 5),\end{aligned}$$

have the same values for $\mu + \tau_1$ and $\mu + \tau_2$, so they give the same distributions for Y_1 and Y_2.

11.6 a. Under the ANOVA assumptions $Y_{ij} = \theta_i + \epsilon_{ij}$, where $\epsilon_{ij} \sim$ independent $\mathrm{n}(0, \sigma^2)$, so $Y_{ij} \sim$ independent $\mathrm{n}(\theta_i, \sigma^2)$. Therefore the sample pdf is

$$\begin{aligned}\prod_{i=1}^k\prod_{j=1}^{n_i}(2\pi\sigma^2)^{-1/2}e^{-\frac{(y_{ij}-\theta_i)^2}{2\sigma^2}} &= (2\pi\sigma^2)^{-\Sigma n_i/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^k\sum_{j=1}^{n_i}(y_{ij}-\theta_i)^2\right\}\\ &= (2\pi\sigma^2)^{-\Sigma n_i/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^k n_i\theta_i^2\right\}\\ &\quad \times \exp\left\{-\frac{1}{2\sigma^2}\sum_i\sum_j y_{ij}^2 + \frac{2}{2\sigma^2}\sum_{i=1}^k \theta_i n_i \bar{Y}_{i\cdot}\right\}.\end{aligned}$$

Therefore, by the Factorization Theorem,

$$\left(\bar{Y}_{1\cdot}, \bar{Y}_{2\cdot}, \dots, \bar{Y}_{k\cdot}, \sum_i\sum_j Y_{ij}^2\right)$$

is jointly sufficient for $(\theta_1, \dots, \theta_k, \sigma^2)$. Since $(\bar{Y}_{1\cdot}, \dots, \bar{Y}_{k\cdot}, S_p^2)$ is a 1-to-1 function of this vector, $(\bar{Y}_{1\cdot}, \dots, \bar{Y}_{k\cdot}, S_p^2)$ is also jointly sufficient.

b. We can write

$$\begin{aligned}&(2\pi\sigma^2)^{-\Sigma n_i/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^k\sum_{j=1}^{n_i}(y_{ij}-\theta_i)^2\right\}\\ &= (2\pi\sigma^2)^{-\Sigma n_i/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^k\sum_{j=1}^{n_i}([y_{ij}-\bar{y}_{i\cdot}]+[\bar{y}_{i\cdot}-\theta_i])^2\right\}\\ &= (2\pi\sigma^2)^{-\Sigma n_i/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^k\sum_{j=1}^{n_i}[y_{ij}-\bar{y}_{i\cdot}]^2\right\}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^k n_i[\bar{y}_{i\cdot}-\theta_i]^2\right\},\end{aligned}$$

so, by the Factorization Theorem, $\bar{Y}_{i\cdot}$, $i = 1, \dots, n$, is independent of $Y_{ij} - \bar{Y}_{i\cdot}$, $j = 1, \dots, n_i$, so S_p^2 is independent of each $\bar{Y}_{i\cdot}$.

c. Just identify $n_i\bar{Y}_{i\cdot}$ with X_i and redefine θ_i as $n_i\theta_i$.

11.7 Let $U_i = \bar{Y}_{i\cdot} - \theta_i$. Then

$$\sum_{i=1}^{k} n_i[(\bar{Y}_{i\cdot} - \bar{\bar{Y}}) - (\theta_i - \bar{\theta})]^2 = \sum_{i=1}^{k} n_i(U_i - \bar{U})^2.$$

The U_i are clearly $\mathrm{n}(0, \sigma^2/n_i)$. For $K = 2$ we have

$$\begin{aligned} S_2^2 &= n_1(U_1 - \bar{U})^2 + n_2(U_2 - \bar{U})^2 \\ &= n_1\left(U_1 - \frac{n_1\bar{U}_1 + n_2\bar{U}_2}{n_1 + n_2}\right)^2 + n_2\left(U_2 - \frac{n_1\bar{U}_1 + n_2\bar{U}_2}{n_1 + n_2}\right)^2 \\ &= (U_1 - U_2)^2\left[n_1\left(\frac{n_2}{n_1 + n_2}\right)^2 + n_2\left(\frac{n_1}{n_1 + n_2}\right)^2\right] \\ &= \frac{(U_1 - U_2)^2}{\frac{1}{n_1} + \frac{1}{n_2}}. \end{aligned}$$

Since $U_1 - U_2 \sim \mathrm{n}(0, \sigma^2(1/n_1 + 1/n_2))$, $S_2^2/\sigma^2 \sim \chi_1^2$. Let $\bar{U}_k$ be the weighted mean of k U_is, and note that

$$\bar{U}_{k+1} = \bar{U}_k + \frac{n_{k+1}}{N_{k+1}}(U_{k+1} - \bar{U}_k),$$

where $N_k = \sum_{j=1}^{k} n_j$. Then

$$\begin{aligned} S_{k+1}^2 &= \sum_{i=1}^{k+1} n_i(U_i - \bar{U}_{k+1})^2 = \sum_{i=1}^{k+1} n_i\left[(U_i - \bar{U}_k) - \frac{n_{k+1}}{N_{k+1}}(U_{k+1} - \bar{U}_k)\right]^2 \\ &= S_k^2 + \frac{n_{k+1}N_k}{N_{k+1}}(U_{k+1} - \bar{U}_k)^2, \end{aligned}$$

where we have expanded the square, noted that the cross-term (summed up to k) is zero, and did a boat-load of algebra. Now since

$$U_{k+1} - \bar{U}_k \sim \mathrm{n}(0, \sigma^2(1/n_{k+1} + 1/N_k)) = \mathrm{n}(0, \sigma^2(N_{k+1}/n_{k+1}N_k)),$$

independent of S_k^2, the rest of the argument is the same as in the proof of Theorem 5.3.1(c).

11.8 Under the oneway ANOVA assumptions, $Y_{ij} \sim$ independent $\mathrm{n}(\theta_i, \sigma^2)$. Therefore

$$\begin{aligned} \bar{Y}_{i\cdot} &\sim \mathrm{n}\left(\theta_i, \sigma^2/n_i\right) \quad (Y_{ij}\text{'s are independent with common } \sigma^2.) \\ a_i\bar{Y}_{i\cdot} &\sim \mathrm{n}\left(a_i\theta_i, a_i^2\sigma^2/n_i\right) \\ \sum_{i=1}^{k} a_i\bar{Y}_{i\cdot} &\sim \mathrm{n}\left(\sum a_i\theta_i, \sigma^2\sum_{i=1}^{k} a_i^2/n_i\right). \end{aligned}$$

All these distributions follow from Corollary 4.6.10.

11.9 a. From Exercise 11.8,

$$T = \sum a_i\bar{Y}_i \sim \mathrm{n}\left(\sum a_i\theta_i, \sigma^2\sum a_i^2\right),$$

and under H_0, $\mathrm{E}T = \delta$. Thus, under H_0,

$$\frac{\sum a_i\bar{Y}_i - \delta}{\sqrt{S_p^2\sum a_i^2}} \sim t_{N-k},$$

where $N = \sum n_i$. Therefore, the test is to reject H_0 if

$$\frac{\sum a_i \bar{Y}_i - \delta}{\sqrt{S_p^2 \sum a_i^2/n_i}} > t_{N-k,\frac{\alpha}{2}}.$$

b. Similarly for $H_0: \sum a_i\theta_i \le \delta$ vs. $H_1: \sum a_i\theta_i > \delta$, we reject H_0 if

$$\frac{\sum a_i \bar{Y}_i - \delta}{\sqrt{S_p^2 \sum a_i^2/n_i}} > t_{N-k,\alpha}.$$

11.10 a. Let H_0^i $i = 1, \ldots, 4$ denote the null hypothesis using contrast a_i, of the form

$$H_0^i: \sum_j a_{ij}\theta_j \ge 0.$$

If H_0^1 is rejected, it indicates that the average of θ_2, θ_3, θ_4, and θ_5 is bigger than θ_1 which is the control mean. If all H_0^i's are rejected, it indicates that $\theta_5 > \theta_i$ for $i = 1, 2, 3, 4$. To see this, suppose H_0^4 and H_0^5 are rejected. This means $\theta_5 > \frac{\theta_5+\theta_4}{2} > \theta_3$; the first inequality is implied by the rejection of H_0^5 and the second inequality is the rejection of H_0^4. A similar argument implies $\theta_5 > \theta_2$ and $\theta_5 > \theta_1$. But, for example, it does not mean that $\theta_4 > \theta_3$ or $\theta_3 > \theta_2$. It also indicates that

$$\frac{1}{2}(\theta_5 + \theta_4) > \theta_3, \quad \frac{1}{3}(\theta_5 + \theta_4 + \theta_3) > \theta_2, \quad \frac{1}{4}(\theta_5 + \theta_4 + \theta_3 + \theta_2) > \theta_1.$$

b. In part a) all of the contrasts are orthogonal. For example,

$$\sum_{i=1}^{5} a_{2i}a_{3i} = \left(0, 1, -\frac{1}{3}, -\frac{1}{3}, -\frac{1}{3}\right)\begin{pmatrix}0\\0\\1\\ -\frac{1}{2}\\ -\frac{1}{2}\end{pmatrix} = -\frac{1}{3} + \frac{1}{6} + \frac{1}{6} = 0,$$

and this holds for all pairs of contrasts. Now, from Lemma 5.4.2,

$$\text{Cov}\left(\sum_i a_{ji}\bar{Y}_{i\cdot}, \sum_i a_{j'i}\bar{Y}_{i\cdot}\right) = \frac{\sigma^2}{n}\sum_i a_{ji}a_{j'i},$$

which is zero because the contrasts are orthogonal. Note that the equal number of observations per treatment is important, since if $n_i \neq n_{i'}$ for some i, i', then

$$\text{Cov}\left(\sum_{i=1}^{k} a_{ji}\bar{Y}_i, \sum_{i=1}^{k} a_{j'i}\bar{Y}_i\right) = \sum_{i=1}^{k} a_{ji}a_{j'i}\frac{\sigma^2}{n_i} = \sigma^2\sum_{i=1}^{k}\frac{a_{ji}a_{j'i}}{n_i} \neq 0.$$

c. This is not a set of orthogonal contrasts because, for example, $a_1 \times a_2 = -1$. However, each contrast can be interpreted meaningfully in the context of the experiment. For example, a_1 tests the effect of potassium alone, while a_5 looks at the effect of adding zinc to potassium.

11.11 This is a direct consequence of Lemma 5.3.3.

11.12 a. This is a special case of (11.2.6) and (11.2.7).

b. From Exercise 5.8(a) We know that

$$s^2 = \frac{1}{k-1}\sum_{i=1}^{k}(\bar{y}_{i\cdot} - \bar{y})^2 = \frac{1}{2k(k-1)}\sum_{i,i'}(\bar{y}_{i\cdot} - \bar{y}_{i'\cdot})^2 .$$

Then

$$\frac{1}{k(k-1)}\sum_{i,i'} t^2_{ii'} = \frac{1}{2k(k-1)}\sum_{i,i'}\frac{(\bar{y}_{i\cdot} - \bar{y}_{i'\cdot})^2}{s_p^2/n} = \sum_{i=1}^{k}\frac{(\bar{y}_{i\cdot} - \bar{\bar{y}})^2}{(k-1)\, s_p^2/n}$$
$$= \frac{\sum_i n(\bar{y}_{i\cdot} - \bar{y})^2/(k-1)}{s_p^2},$$

which is distributed as $F_{k-1,N-k}$ under $H_0: \theta_1 = \cdots = \theta_k$. Note that

$$\sum_{i,i'} t^2_{ii'} = \sum_{i=1}^{k}\sum_{i'=1}^{k} t^2_{ii'},$$

therefore $t^2_{ii'}$ and $t^2_{i'i}$ are both included, which is why the divisor is $k(k-1)$, not $\frac{k(k-1)}{2} = \binom{k}{2}$. Also, to use the result of Example 5.9(a), we treated each mean $\bar{Y}_{i\cdot}$ as an observation, with overall mean $\bar{\bar{Y}}$. This is true for equal sample sizes.

11.13 a.

$$L(\theta|y) = \left(\frac{1}{2\pi\sigma^2}\right)^{Nk/2} e^{-\frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij}-\theta_i)^2/\sigma^2} .$$

Note that

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij}-\theta_i)^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{i\cdot})^2 + \sum_{i=1}^{k} n_i(\bar{y}_{i\cdot}-\theta_i)^2$$
$$= SSW + \sum_{i=1}^{k} n_i(\bar{y}_{i\cdot}-\theta_i)^2,$$

and the LRT statistic is

$$\lambda = (\hat{\tau}^2/\hat{\tau}_0^2)^{Nk/2}$$

where

$$\hat{\tau}^2 = SSW \quad \text{and} \quad \hat{\tau}_0^2 = SSW + \sum_i n_i(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2 = SSW + SSB.$$

Thus $\lambda < k$ if and only if SSB/SSW is large, which is equivalent to the F test.

b. The error probabilities of the test are a function of the θ_is only through $\eta = \sum \theta_i^2$. The distribution of F is that of a ratio of chi squared random variables, with the numerator being noncentral (dependent on η). Thus the Type II error is given by

$$P(F > k|\eta) = P\left(\frac{\chi^2_{k-1}(\eta)/(k-1)}{\chi^2_{N-k}/(N-k)} > k\right) \geq P\left(\frac{\chi^2_{k-1}(0)/(k-1)}{\chi^2_{N-k}/(N-k)} > k\right) = \alpha,$$

where the inequality follows from the fact that the noncentral chi squared is stochastically increasing in the noncentrality parameter.

11.14 Let $X_i \sim \mathrm{n}(\theta_i, \sigma^2)$. Then from Exercise 11.11

$$\mathrm{Cov}\left(\sum_i \frac{a_i}{\sqrt{c_i}} X_i, \sum_i \sqrt{c_i} v_i X_i\right) = \sigma^2 \sum a_i v_i$$

$$\mathrm{Var}\left(\sum_i \frac{a_i}{\sqrt{c_i}} X_i\right) = \sigma^2 \sum \frac{a_i^2}{c_i}, \qquad \mathrm{Var}\left(\sum_i \sqrt{c_i} v_i X_i\right) = \sigma^2 \sum c_i v_i^2,$$

and the Cauchy-Schwarz inequality gives

$$\left(\sum a_i v_i\right) \Big/ \left(\sum a_i^2/c_i\right) \le \sum c_i v_i^2.$$

If $a_i = c_i v_i$ this is an equality, hence the LHS is maximized. The simultaneous statement is equivalent to

$$\frac{\left(\sum_{i=1}^k a_i(\bar{y}_{i\cdot} - \theta_i)\right)^2}{s_p^2 \sum_{i=1}^k a_i^2/n} \le M \quad \text{for all } a_1, \ldots, a_k,$$

and the LHS is maximized by $a_i = n_i(\bar{y}_{i\cdot} - \theta_i)$. This produces the F statistic.

11.15 a. Since $t_v^2 = F_{1,v}$, it follows from Exercise 5.19(b) that for $k \ge 2$

$$P[(k-1)F_{k-1,v} \ge a] \ge P(t_v^2 \ge a).$$

So if $a = t^2_{v,\alpha/2}$, the F probability is greater than α, and thus the α-level cutoff for the F must be greater than $t^2_{v,\alpha/2}$.

b. The only difference in the intervals is the cutoff point, so the Scheff´e intervals are wider.

c. Both sets of intervals have nominal level $1-\alpha$, but since the Scheff´e intervals are wider, tests based on them have a smaller rejection region. In fact, the rejection region is contained in the t rejection region. So the t is more powerful.

11.16 a. If $\theta_i = \theta_j$ for all i, j, then $\theta_i - \theta_j = 0$ for all i, j, and the converse is also true.

b. $H_0 : \boldsymbol{\theta} \in \cap_{ij}\Theta_{ij}$ and $H_1 : \boldsymbol{\theta} \in \cup_{ij}(\Theta_{ij})^c$.

11.17 a. If all of the means are equal, the Scheff´e test will only reject α of the time, so the t tests will be done only α of the time. The experimentwise error rate is preserved.

b. This follows from the fact that the t tests use a smaller cutoff point, so there can be rejection using the t test but no rejection using Scheff´e. Since Scheff´e has experimentwise level α, the t test has experimentwise error greater than α.

c. The pooled standard deviation is 2.358, and the means and t statistics are

Mean			t statistic		
Low	Medium	High	Med-Low	High-Med	High-Low
3.51.	9.27	24.93	3.86	10.49	14.36

The t statistics all have 12 degrees of freedom and, for example, $t_{12,.01} = 2.68$, so all of the tests reject and we conclude that the means are all significantly different.

11.18 a.

$$\begin{aligned} P(Y > a \mid Y > b) &= P(Y > a, Y > b)/P(Y > b) \\ &= P(Y > a)/P(Y > b) && (a > b) \\ &> P(Y > a). && (P(Y > b) < 1) \end{aligned}$$

b. If a is a cutoff point then we would declare significance if $Y > a$. But if we only check if Y is significant because we see a big Y $(Y > b)$, the proper significance level is $P(Y > a \mid Y > b)$, which will show less significance than $P(Y > a)$.

11.19 a. The marginal distributions of the Y_i are somewhat straightforward to derive. As $X_{i+1} \sim \text{gamma}(\lambda_{i+1}, 1)$ and, independently, $\sum_{j=1}^{i} X_j \sim \text{gamma}(\sum_{j=1}^{i} \lambda_j, 1)$ (Example 4.6.8), we only need to derive the distribution of the ratio of two independent gammas. Let $X \sim \text{gamma}(\lambda_1, 1)$ and $Y \sim \text{gamma}(\lambda_2, 1)$. Make the transformation

$$u = x/y, \quad v = y \quad \Rightarrow \quad x = uv, \quad y = v,$$

with Jacobian v. The density of (U, V) is

$$f(u, v) = \frac{1}{\Gamma(\lambda_1)\Gamma(\lambda_2)}(uv)^{\lambda_1 - 1} v^{\lambda_2 - 1} v e^{-uv} e^{-v} = \frac{u^{\lambda_1 - 1}}{\Gamma(\lambda_1)\Gamma(\lambda_2)} v^{\lambda_1 + \lambda_2 - 1} e^{-v(1+u)}.$$

To get the density of U, integrate with respect to v. Note that we have the kernel of a $\text{gamma}(\lambda_1 + \lambda_2, 1/(1 + u))$, which yields

$$f(u) = \frac{\Gamma(\lambda_1 + \lambda_2)}{\Gamma(\lambda_1)\Gamma(\lambda_2)} \frac{u^{\lambda_1 - 1}}{(1 + u)^{\lambda_1 + \lambda_2 - 1}}.$$

The joint distribution is a nightmare. We have to make a multivariate change of variable. This is made a bit more palatable if we do it in two steps. First transform

$$W_1 = X_1, \quad W_2 = X_1 + X_2, \quad W_3 = X_1 + X_2 + X_3, \quad \ldots, \quad W_n = X_1 + X_2 + \cdots + X_n,$$

with

$$X_1 = W_1, \quad X_2 = W_2 - W_1, \quad X_3 = W_3 - W_2, \quad \ldots \quad X_n = W_n - W_{n-1},$$

and Jacobian 1. The joint density of the W_i is

$$f(w_1, w_2, \ldots, w_n) = \prod_{i=1}^{n} \frac{1}{\Gamma(\lambda_i)} (w_i - w_{i-1})^{\lambda_i - 1} e^{-w_n}, \quad w_1 \le w_2 \le \cdots \le w_n,$$

where we set $w_0 = 0$ and note that the exponent telescopes. Next note that

$$y_1 = \frac{w_2 - w_1}{w_1}, \quad y_2 = \frac{w_3 - w_2}{w_2}, \quad \ldots \quad y_{n-1} = \frac{w_n - w_{n-1}}{w_{n-1}}, \quad y_n = w_n,$$

with

$$w_i = \frac{y_n}{\prod_{j=i}^{n-1}(1 + y_j)}, \quad i = 1, \ldots, n-1, \quad w_n = y_n.$$

Since each w_i only involves y_j with $j \ge i$, the Jacobian matrix is triangular and the determinant is the product of the diagonal elements. We have

$$\frac{dw_i}{dy_i} = -\frac{y_n}{(1 + y_i)\prod_{j=i}^{n-1}(1 + y_j)}, \quad i = 1, \ldots, n-1, \quad \frac{dw_n}{dy_n} = 1,$$

and

$$\begin{aligned} f(y_1, y_2, \ldots, y_n) &= \frac{1}{\Gamma(\lambda_1)} \left(\frac{y_n}{\prod_{j=1}^{n-1}(1 + y_j)} \right)^{\lambda_1 - 1} \\ &\times \prod_{i=2}^{n-1} \frac{1}{\Gamma(\lambda_i)} \left(\frac{y_n}{\prod_{j=i}^{n-1}(1 + y_j)} - \frac{y_n}{\prod_{j=i-1}^{n-1}(1 + y_j)} \right)^{\lambda_i - 1} e^{-y_n} \\ &\times \prod_{i=1}^{n-1} \frac{y_n}{(1 + y_i)\prod_{j=i}^{n-1}(1 + y_j)}. \end{aligned}$$

Factor out the terms with y_n and do some algebra on the middle term to get

$$
\begin{aligned}
f(y_1, y_2, \ldots, y_n) &= y_n^{\sum_i \lambda_i - 1} e^{-y_n} \frac{1}{\Gamma(\lambda_1)} \left(\frac{1}{\prod_{j=1}^{n-1}(1+y_j)} \right)^{\lambda_1 - 1} \\
&\quad \times \prod_{i=2}^{n-1} \frac{1}{\Gamma(\lambda_i)} \left(\frac{y_{i-1}}{1+y_{i-1}} \frac{1}{\prod_{j=i}^{n-1}(1+y_j)} \right)^{\lambda_i - 1} \\
&\quad \times \prod_{i=1}^{n-1} \frac{1}{(1+y_i) \prod_{j=i}^{n-1}(1+y_j)}.
\end{aligned}
$$

We see that Y_n is independent of the other Y_i (and has a gamma distribution), but there does not seem to be any other obvious conclusion to draw from this density.

b. The Y_i are related to the F distribution in the ANOVA. For example, as long as the sum of the λ_i are integers,

$$
Y_i = \frac{X_{i+1}}{\sum_{j=1}^{i} X_j} = \frac{2X_{i+1}}{2\sum_{j=1}^{i} X_j} = \frac{\chi^2_{\lambda_{i+1}}}{\chi^2_{\sum_{j=1}^{i}\lambda_j}} \sim F_{\lambda_{i+1}, \sum_{j=1}^{i}\lambda_j}.
$$

Note that the F density makes sense even if the λ_i are not integers.

11.21 a.

$$
\begin{aligned}
\text{Grand mean } \bar{y}_{\cdot\cdot} &= \frac{188.54}{15} = 12.57 \\
\text{Total sum of squares} &= \sum_{i=1}^{3}\sum_{j=1}^{5} (y_{ij} - \bar{y}_{\cdot\cdot})^2 = 1295.01. \\
\text{Within SS} &= \sum_{1}^{3}\sum_{1}^{5} (y_{ij} - \bar{y}_{i\cdot})^2 \\
&= \sum_{1}^{5} (y_{1j} - 3.508)^2 + \sum_{1}^{5} (y_{2j} - 9.274)^2 + \sum_{1}^{5} (y_{3j} - 24.926)^2 \\
&= 1.089 + 2.189 + 63.459 = 66.74 \\
\text{Between SS} &= 5 \sum_{1}^{3} \left(y_{ij} - \bar{y}_{i\cdot} \right)^2 \\
&= 5(82.120 + 10.864 + 152.671) = 245.65 \cdot 5 = 1228.25.
\end{aligned}
$$

ANOVA table:

Source	df	SS	MS	F
Treatment	2	1228.25	614.125	110.42
Within	12	66.74	5.562	
Total	14	1294.99		

Note that the total SS here is different from above – round off error is to blame. Also, $F_{2,12} = 110.42$ is highly significant.

b. Completing the proof of (11.2.4), we have

$$
\sum_{i=1}^{k}\sum_{j=1}^{n_i} (y_{ij} - \bar{\bar{y}})^2 = \sum_{i=1}^{k}\sum_{j=1}^{n_i} \left((y_{ij} - \bar{y}_{i\cdot}) + (\bar{y}_{i\cdot} - \bar{\bar{y}}) \right)^2
$$