

1. (a)

$$X = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 \end{bmatrix}$$

If you only wrote out three rows, you're describing a study with 3 subjects, not this study with 6 subjects.

(b) In terms of model parameters, the four cell means and the column marginal means are:

Study time	Absent	Present
1 min	$\mu + \alpha_1 + \beta_1$	$\mu + \alpha_1 + \beta_2$
5 min	$\mu + \alpha_2 + \beta_1$	$\mu + \alpha_1 + \beta_2$
ave.	$\mu + (\alpha_1 + \alpha_2)/2 + \beta_1$	$\mu + (\alpha_1 + \alpha_2)/2 + \beta_2$

Hence the main effect of refresher is the column mean for absent - column mean for present =  $\beta_1 - \beta_2$ .

Yes, this is estimable. One possible linear combination of rows is obs. 2 - obs. 3, which has the desired expected value.

Note: A few of you used a different definition of the average difference between present and absent. I accepted that. But, please realize that is not the customary definition of main effect.

(c) Yes, this is estimable. One possible linear combination of rows is obs. 5 - obs. 2 + obs. 3., i.e.  $\mu_{5A} - (\beta_1 - \beta_2)$ .

2. (a)  $15 \times 15$ .

You can work this out in at least two ways:

$P_X = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ ,  $\mathbf{X}$  is  $15 \times 8$ , and work out the dimension of the product.

$P_X$  maps a vector of 15 observations into a vector of 15 expected values. Hence it has to have 15 columns for the matrix multiplication to be conformable. And it has to have 15 rows because the result is a vector of 15 observations.

(b) error df =  $9 = n - \text{rank}\mathbf{X} = 15 - 6$ .

A lot of folks said  $\text{rank}\mathbf{X} = 8$ . It has 8 columns but two are linearly dependent on the others. If you remembered the assigned readings on eigen decompositions, you remembered that  $\text{rank}(\mathbf{X})$  is the number of non-zero eigenvalues. Or, you might have remembered that  $\text{rank}(P_X) = \text{trace}(P_X) = \text{trace}(UDU') = \text{trace}(DU'U) = \text{trace}(D)$ . That's 6.

(c) 0.6

(d) 0.352

(e)  $T = \frac{0.6}{\sqrt{0.352}} = 0.66$ . Under  $H_0$  this has a central T distribution with 9 d.f.

I accepted answers that omitted “central”. If I wrote ok, it’s because you reported the same d.f. as in part b. I try not to deduct points twice for the same mistake.

(f) either  $ncp = 1.11$  if you report the T non-centrality or  $ncp = 1.23$  if you report the F non-centrality.

I tell you that  $\beta_1 = 1$ , so  $\mathbf{C}\beta - d = 1 - 0 = 1$ , and that the 15 subjects have the same configuration of X’s as the obs. data, so  $\text{Var } \hat{\beta}_1 = 0.352 \times 2.3$ .

Hence, for the F test,  $ncp = 1 \times 0.352^{-1} \times 1/2.3 = 1.23$ . The T value is  $\sqrt{1.23}$ .

(g)  $0.31 = 0.54 - 0.23$

(h)  $0.352 \times 2.3 = 2.43$ .

$\text{Var } \hat{\beta}_3 - \hat{\beta}_5 = \text{Var } \hat{\beta}_3 + \text{Var } \hat{\beta}_5 - 2 \text{Cov } \hat{\beta}_3, \hat{\beta}_5 = (0.271 + 0.295 - 2 \times (-0.246)) \times 2.3 = 1.058 \times 2.3 = 2.43$

Quite a few folks forgot to include the covariance.

(i) The test of  $H_0: \beta_1 = 0$  has the highest power.

All three tests are t-tests with the same difference and the same error d.f. The test with the highest power is the one with the largest  $ncp$ , which is the one with the smallest variance for the associated estimate. That is  $\hat{\beta}_1$ .

3. (a)  $y_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk}$ ,  $\epsilon_{ijk} \sim (0, \sigma^2 x_j)$ , where  $i \in \{1, 2\}$  indexes the animal,  $j \in \{1, 2, 3, 4\}$  indexes the aging time, and  $k \in \{1, 2\}$  indexes the steak within animal and aging time.

$\alpha_i$  describes differences among animals,

$\beta_j$  describes differences among aging times

$\epsilon_{ijk}$  describes variation among steaks from the same animal and aged for the same amount of time.

$x_j$  is the number of days of aging

I accepted  $y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \epsilon_{ijk}$ ,  $\epsilon_{ijk} \sim (0, \sigma^2 x_j)$  because it is not clear whether interactions should be separated from the error variation.

A lot of folks described the error as  $N(0, \sigma^2 x_j)$ . The problem specifies the variance; it does not tell you anything about the distribution. If I crossed out the N and deducted two points, that’s why.

(b) Aitken model

(c)

$$\Sigma = \begin{bmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & 3\sigma^2 & 0 & 0 \\ 0 & 0 & 7\sigma^2 & 0 \\ 0 & 0 & 0 & 14\sigma^2 \end{bmatrix}$$

The problem tells you the variances of the errors are  $\sigma^2 x_j$  and that errors are independent. That specifies this  $\Sigma$ .

This is the answer even if you considered animals to be random blocks; the problem asks for the VC matrix of the **errors**, not the observations.

- (d) The model  $X_0 + B_1 + X_1 + X_2$  vs the model  $X_0 + B_1 + X_1$ .

This is the definition of sequential SS

(e)  $F = \frac{1.115+0.121+0.068)/3}{0.6084/11} = 7.85$ .

This is the comparison of the models  $X_0 + B_1 + X_1 + X_2 + X_3$  vs the model  $X_0 + B_1$ . Since the  $X_1$ ,  $X_2$ , and  $X_3$  terms are consecutive, you can add the sequential SS. The appropriate error MS is that from the full model.

Some folks answered 9.41, which is what SAS reports as the model SS. This is a comparison of  $X_0 + B_1 + X_1 + X_2 + X_3$  vs  $X_0$ , so it includes blocks.

- (f) numerator 3 df, denominator 11 df

- (g) The error variance for an observation on day 3 is  $3\sigma^2$ , so the estimate of that is  $3MSE = 3 \times 0.0553 = 0.166$ .

A lot of folks answered with something including elements of  $(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}$ . I wasn't always sure what you were trying to do, but in general  $(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}$  answers questions about the standard errors of  $\hat{\beta}$ . My question asked about the error variance, not the precision of an estimate.

- (h)  $X_0 + B_1 + X_1 + X_2 + X_3$  vs  $X_0 + B_1 + X_1 + X_3$

i.e., the consequence of removing  $X_2$  from a model with all terms.

- (i) 0.1215, i.e., the same as the sequential SS.

This is because all effects are orthogonal (look at  $\mathbf{X}\Sigma^{-1}\mathbf{X}$ , which is diagonal). So the sequential SS are the same as the partial SS.

I accepted "because the design is balanced" as sufficient explanation. In fact, in Aitken models that isn't sufficient. The estimates have to be orthogonal.