1. Eight individuals were given a drug to measure the effect of a sopori_c effect on alertness. They were measured 5, 10, 15, 20, and 25 minutes afterward. At each time point, each person's response time (in hundredths of a second) to a stimulus was recorded. Each person's sex was also recorded for consideration. Output for the code below is available in \ex2 drug.pdf".

a)
When we are testing for interaction we are testing the hypothesis that these lines segments are parallel to one another. From the plot we see that the lines are approximately parallel to each other and based on that we can say that there isn't significant interaction between the gender and time.
This means the response rates for each gender are equal for all times.
For time points 10, 15, 20 and 25, the mean response rates are higher for males. However, at time point 5 female is higher.

b) Use the output to carry out the test of interaction. State the test statistic and conclusion with alpha = .05
We have:

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall sex Effect on the Variables Defined by the M Matrix Transformation
H = Type III SSCP Matrix for sex
E = Error SSCP Matrix

S=1 M=1 N=0.5

Statistic	Value	F Value	Num DF	Den DF
Wilks' Lambda	0.86482775	0.12	4	3
Pillai's Trace	0.13517225	0.12	4	3
Hotelling-Lawley Trace	0.15629962	0.12	4	3
Roy's Greatest Root	0.15629962	0.12	4	3

Here we get a Wilk's Lambda of 0.865 with a supporting F-value of 0.12 with 4 and 3 d.f.
Fcritical (4,3,.05) = 9.117182253

The null hypothesis is H0 : no interaction between time and sex. The Wilk's Lambda test statistic is 0.865 with 4 and 3 degrees of freedom. We have F-value < Fcritical and the p-value = 0.9673 > .05. We fail to reject the null hypothesis that sex and time do not interact.

c) Is there an effect due to sex? Use = :05 to conduct the appropriate test, based on the output. Also, comment on the appropriateness of this test in light of your results above in the test for interaction.

We have:

MANOVA Test Criteria and Exact F Statistics for the Hypothesis of No Overall sex Effect on the Variables Defined by the M Matrix Transformation H = Type III SSCP Matrix for sex E = Error SSCP Matrix S=1 M=-0.5 N=2				
Statistic	**Value**	**F Value**	**Num DF**	**Den DF**
Wilks' Lambda	0.99092705	0.05	1	6
Pillai's Trace	0.00907295	0.05	1	6
Hotelling-Lawley Trace	0.00915602	0.05	1	6
Roy's Greatest Root	0.00915602	0.05	1	6

Here we get a Wilk's Lambda of 0.991 with a supporting F-value of 0.05 with 1 and 6 d.f.
Fcritical (1,6,.05) = 5.987

This indicates that there is no significant main effect of sex. That is that the mean response of our different time variables doesn't differ significantly among sexes.

Conclusion: Sex doesn't have a significant effect on the response rate over the different time points ($\Lambda = 0.991$; $F = 0.05$; d. f. = 1, 6; $p = 0.8225$).

Appropriateness: The test results seem appropriate when you also look at the profile plot. The mean response of both the sexes is reducing at a similar rate. The variation seems to be on account of time rather than the difference in the sexes of the individuals.

2. Fish

a) The variables are: aroma score, flavor score, texture score, moisture score.
We have 3 groups with 3 different preparation methods. We have 12 fishes in each group.
Y_{ijk} = Observation for variable k (one of the aroma score, flavor score, texture score, moisture score) from subject j (one of 12 fishes) in group i (one of 3 methods)

Here we are interested in testing the null hypothesis that group mean vectors are all equal to one another. Mathematically this is expressed as:

$H_0: \mu_1 = \mu_2 = \mu_3$

The alternative hypothesis being: $H_a: \mu_{ik} \neq \mu_{jk}$ for at least one $i \neq j$ and at least one variable k

This says that the null hypothesis is false if at least one pair of treatments is different on at least one variable.

b) We see here that the preparation method has an impact on the mean score for almost all of the variables. Under the assumption of no method effect, we shouldn't see as much variation in the scores for all of the variables. Even if one of the variables has significant variation we will fail to reject the null hypothesis of no method effect.

c) This requires that we start with the null hypothesis:

$H0:\mu1=\mu2=\mu3$

The alternative hypothesis being: $Ha:\mu ik \neq \mu jk$ for at least one $i \neq j$ and at least one variable k

We have from the output:

DV	F	SAS *p*-value	Df
Aroma	**1.29**	**0.2880**	**2, 33**
Flavor	9.38	0.0006	**2, 33**
Texture	3.39	0.0460	**2, 33**
Moisture	1.27	0.2954	**2, 33**

Analysis of Individual Chemical Elements – Naïve approach

$F(2, 33, 0.05) = 3.2849$
No we see that F-value is > Fcritical for Flavor and Texture dimensions. Also for these variables $p < 0.05$ and therefore we see that for Flavor and Texture have significant results

Analysis of Individual Chemical Elements – Bonferroni correction
Here, p = 4 variables, g = 3 groups, and a total of N = 36 observations. So, for an $\alpha = 0.05$ level test, we reject

$H0:\mu 1k = \mu 2k = \cdots = \mu gk$ if $F > F(g-1, N-g, \alpha/p) = F(2, 33, 0.05/4) = 5.019$

Since only the F-value for Flavor exceeds the critical value of 5.019, or equivalently, since the SAS p-value of flavor falls below 0.0125 (= 0.05/4), we can see that only flavor has significant results at the 0.05 level under the Bonferroni correction.

Conclusion: We have enough evidence to reject the null hypothesis since flavor is significantly different for the 3 different preparation methods.

d) The assumption is that data from all groups have common variance-covariance matrix Σ.

We will rely on the Bartlett's test here:

Test of Homogeneity of Within Covariance Matrices

Chi-Square	DF	Pr > ChiSq
12.473068	20	0.8988

Since the Chi-Square value is not significant at the 0.1 level, a pooled covariance matrix will be used in the discriminant function.

We find no statistically significant evidence against the null hypothesis that the variance-covariance matrices are homogeneous ($L' = 12.47$; d.f. = 20; $p = 0.8988$).

We can reasonably assume that all groups have common variance-covariance matrix Σ

3. Discriminant Analysis

   a) The variables are appropriate as discriminators of admittance from the plot. There is a clear classification line that is indicated by the plot. Further the plot illustrates somewhat elliptical scatters for each group individually, suggesting bivariate normality for each group individually, but the orientations of the ellipses suggest different covariance matrices.
   b) Using gpa and gmat, the APER = 1/31+1/54 = 0.0254

      $\hat{d}_1(\mathbf{y}) = -134.68959 + 48.54 \text{ gpa} + 0.183 \text{ gmat}$

      $\hat{d}_2(\mathbf{y}) = -86.30887 + 38.88 \text{ gpa} + 0.146 \text{ gmat}$

      Misclassification is .0254 and correct is = 0.9746

   c) Now, consider a student with the following:

GPA and GMAT score of 2.9 and 550

$$\hat{d}_1(\mathbf{y}) = -134.68959 + 48.54 * 2.9 + 0.183 * 550 = 106.72641$$

$$\hat{d}_2(\mathbf{y}) = -86.30887 + 38.88 * 2.9 + 0.146 * 550 = 106.74313$$

Then the linear score function is obtained by adding in a log of one half, here for group 1 (yes)

$\hat{s}^L_a(x)=\hat{d}^L_a(x) + \log \hat{p}_a = 106.72641 + \log(0.5) = 106.0333$

and then for group 2 (no):

$\hat{s}^L_b(x)=\hat{d}^L_b(x) + \log \hat{p}_b = 106.74313 + \log(0.5) = 106.05$

Conclusion

According to the classification rule the student is classified into the group 2 (no) since that has the highest linear discriminant function.

Posterior probability of accepting = EXP(106.0333) / (EXP(106.0333) + EXP(106.05)) = 0.495825097

Posterior probability of group 2 or rejecting = EXP(106.05) / (EXP(106.0333) + EXP(106.05)) = 0.504174903

d) We make the following assumptions:
   - assume equal priors (equal probability)
   - We assume that in population πi the probability density function of x is multivariate normal with mean vector μi and variance-covariance matrix Σ
   - The subjects are independently sampled.