

1a. Name the assumptions that underlie the accurate computation of p-values from F and t distributions. [4]

1. residuals homogeneous (fixed variance, homoscedastic)
2. $\sum \text{residuals} = 0$
3. residuals independent
4. residuals normal

1b List one assumption and state how you would check the assumption. [2]

1. plot residuals versus fits, check for violation (unequal vertical distribution)
2. $\sum \text{residuals} = 0$ for fitted values produced by statistical packages.
3. plot residuals against neighboring value, check for no trend (no correlation)
4. examine histogram, examine normal plot for deviations from straight line

2. A geneticist analyzes survival (in days) of the fruit fly *Drosophila subobscura* in relation to lifetime production of eggs (Prod) and egg size as measured by length (*L*).

2a. What units will the mean survival have ? days [1]

Complete the ANOVA table. [9]

Source	df	SS	MS	F
Prod	<u>1</u>	3.340	<u>3.34</u>	<u>8.48</u>
<i>L</i>	<u>1</u>	21.842	<u>21.842</u>	<u>55.45</u>
Error	<u>22</u>	8.665	<u>0.394</u>	
Total	24	<u>33.85</u>		

2b. What is the variance in survival of all 25 flies? $33.85/24 = 1.41 \text{ days}^2$ [1]

2c. If egg size *L* is omitted from the analysis of survival in relation to egg length, state what happens to each of the following (increase or decrease?). [3]

Error degrees of freedom increase

MS error increase

F- ratio for Prod decrease