1. (a) \( \mu_0 = 26 \) bird species  
(b) \( \alpha = 0.10 \)  
(c) \( \mu = 28.7 \) species approximately will be detected with 80% power, so it makes a difference of 2.7 species.  
(d) The dotted curve, because more plots gives more power (if the level \( \alpha \) remains the same)

(a) Slope: \( b_1 = 1192.4/1105.6 = 1.0785 \).  
Intercept: \( b_0 = -36.366 \).

(b) Using Anova:  
\[
\begin{array}{l|cccc}
\text{df} & \text{SS} & \text{MS} & F & \text{p-value} \\
1 & 1286 & 1286 & 10.85 & .005 < p < .01 \\
13 & 1630.92 & 125.45 & & \\
14 & 2916.93 & & & \\
\end{array}
\]

or using a t-test: \( s^2 = 1630.92/(15 - 2) = 125.45 \) then SE of \( b_1 \) is \( \sqrt{125.45/1105.6} = .337 \) so \( t = 3.20 \) and we get \(.002 < p < .01 \). Either way, there is strong evidence of a positive relationship between percentage of normal spermatozoa and percentage of male offsprings.

(c) Prediction \( y = 33.7 \) with standard error \( \sqrt{\frac{125.45}{15} + \frac{(65-79.4)^2}{1105.6}} + 1 = 12.54 \). For 95% confidence: \( 33.7 \pm 2.160 \times 12.54 \) i.e. (6.61, 60.8) percent male offsprings.

3. (a) \( df = 2 \) so \(.025 < p < .05 \). There is moderate evidence that a fish’s environment preference is linked to its color.  
(b) For this cell the expected value is 34.45 and the contribution to the \( X^2 \) value is 0.0698. 
(c) Assumptions: independence seems okay. Expected counts > 5 is also met, i.e. the sample size was large enough for the chi-square test to yield an accurate p-value.

4. Log-transformed data. Reasons: the normal scores plot looks more straight (normality of residuals) and the samples look to have more similar variances.

5. (a) \( \text{SSErr} = 6 \times .84^2 + 6 \times 1.30^2 + 6 \times 1.15^2 + 6 \times 1.26^2 = 31.83 \).

We can obtain SSTreatment by first calculating the grand mean \( \frac{7 \times 1.1 + 7 \times 2.6 + \cdots + 7 \times 10.3}{28} = 5.2 \) then SSTreatment = \( 7 \times (1.1 - 5.2)^2 + \cdots + 7 \times (10.3 - 5.2)^2 = 365 \).

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>3</td>
<td>365</td>
<td>121.66</td>
<td>91.7</td>
</tr>
<tr>
<td>Error</td>
<td>24</td>
<td>31.83</td>
<td>1.326</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>27</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The p-value is much < .001. We strongly reject the null hypothesis. There is strong evidence that the 4 techniques have different mean survival times.

(b) \( s^2 = 1.326 \) associated with \( df = 24 \). For LSD we get \( d_L = 2.797 \sqrt{1.326 \times 2/7} = 1.721 \). \( d_Q = 4.91 \sqrt{1.326 \times 1/7} = 2.317 \) for Tukey. For Bonferroni, we need to know \( t_{.005/3,24} = t_{.001,24} = 3.467 \). We can round down (not up). We get \( d_B = 3.4 \sqrt{1.326 \times 2/7} = 2.093 \). Significant differences:

<table>
<thead>
<tr>
<th>Diff</th>
<th>LSD</th>
<th>Tukey</th>
<th>Bonf.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bar{y}_B - \bar{y}_A )</td>
<td>1.5</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>( \bar{y}_D - \bar{y}_A )</td>
<td>9.2</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>( \bar{y}_C - \bar{y}_B )</td>
<td>4.2</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>

6. (a) 4 regions. We read \( df = 78 \) in the R output, so \( n_1 + n_2 - 2 = 78 \). Since the 2 samples have same size, we have \( n_1 + n_2 = 80 \) and \( n_1 = n_2 = 40 \). Each region was made of 10 forest plots and 10 agriculture plots, so there were 4 regions.

(b) Independence: not met because plots were nested within regions. The plots from the same region are possibly more similar to each other than to plots from other regions.

Normality: seems met, because both normal scores plots are very linear.

Equal variance across the 2 samples: seems met, from the first plot.

7. (a) False. Statistics uses samples, probability does not.
(b) False. The equality may not hold if the random variables are not independent.
(c) False. It’s from a sum of squares.
(d) True. CLT
(e) True. The T-distribution is symmetric around 0, so its mean is 0.
(f) False. The null hypothesis is either true or false. Whether it is true is not random.
(g) False. \( P(12 \leq \mu \leq 32) \) is either 0 or 1.
(h) True. That’s the definition of a quantile.