1. (a) $0.9158 \pm t_{23}(0.975) \times 0.0438 = 0.9158 \pm 0.0438 = (0.825, 1.006)$.

(b) 99% CI for $\sigma^2 = \left( \frac{23 \times 0.2144^2}{\chi^2_{23}(0.995)}, \frac{23 \times 0.2144^2}{\chi^2_{23}(0.005)} \right) = \left( \frac{1.057}{441.18}, \frac{1.057}{9.266} \right) = (0.02392, 0.1141); (0.02392, 0.1141)^{1/2} = (0.155, 0.338).

(c) 95% CI includes 1, so accept hypothesis. $H_0 : \mu = 1; H_1 : \mu \neq 1$.

(d) Rejection region is $\bar{X}$ outside the range $1 \pm 2.069 \times 0.0438 = (0.909, 1.091)$.

(e) $Z = \frac{(\bar{X} - 1.15)}{0.2/\sqrt{24}} = N(0, 1)$. $Pr(\bar{X} < 0.909|\mu = 1.15) + Pr(\bar{X} > 1.091|\mu = 1.15) = Pr(Z < \frac{0.909 - 1.15}{0.2/\sqrt{24}}) + Pr(Z > \frac{1.091 - 1.15}{0.2/\sqrt{24}}) = 0 + 0.9258 = 0.9258$.

(f) Pooled $s^2 = (23 \times 0.2144^2 + 7 \times 0.2428^2)/30 = 0.0490; \sqrt{0.0490} = 0.221$. $T = \frac{0.9763 - 0.9158}{0.221/\sqrt{24}+1/8} = 0.671 < t_{30}(0.95) = 1.697$, so accept hypothesis. $0.671 \approx t_{30}(0.75)$, so $P \approx 0.5$.

2. (a) plot: see attached; negative relationship between $y$ and $x$; two influential points with large $x$.

(b) $r = -0.400 = -0.632$.

(c) $6.171 - 0.002083 \times 350 = 5.44; \bar{x} = 281.9$ (by calculator), so s.e. of expected value = $\sqrt{0.1582^2/12 + (350 - 281.9)^2 \times 0.0008076^2} = 0.0715$; 95% CI = $5.44 \pm t_{10}(0.975) \times 0.0715 = 5.44 \pm 2.228 \times 0.0715 = (5.28, 5.60)$.

(d) It is the estimated sweetness of juice when no pectin is present; this is not useful because pectin always seems to be present in substantial quantities.

(e) res SS = $s^2 \times$ res df = 0.1582 $\times 10 = 0.2503$; total SS = res SS/(1 - $R^2$) = 0.2503/0.600 = 0.4172 (or calculate total SS = $(n - 1)s^2_{alt}$ by calculator); adj $R^2 = 1 - \frac{0.2503/10}{0.4172/11} = 0.340$. 
3.
(a) Source DF SS MS F
machine 5 1722.3 344.5 4.12
Error 18 1503.7 83.54
Total 23 3226.0

4.12 > F_{5,18}(0.95) = 2.77, so reject hypothesis at the 0.05 level.

(b) $176.75 \pm t_{18}(0.975) \times 9.14/\sqrt{4} = 176.75 \pm 2.101 \times 4.57 = (167.15, 186.35)$.

(c) $185.00 - 176.75 \pm Q_{6,18}(0.95)/\sqrt{2} \times 9.14\sqrt{1/4 + 1/4} = 8.25 \pm 4.505/\sqrt{2} \times 6.463 = 8.25 \pm 20.59 = (-12.34, 28.84)$.

(d) $\psi = \frac{\mu_2 + \mu_3}{2} - \frac{\mu_4 + \mu_5 + \mu_6}{3}; \tilde{\psi} = \frac{(171.75 + 176.75)\sqrt{2} - (185.00 + 184.25 + 189.00)}{3} = -11.83$; $\text{s.e.}(\tilde{\psi}) = 9.14\sqrt{\frac{(1/2)^2}{4} + \frac{(1/2)^2}{4} + \frac{(-1/3)^2}{4} + \frac{(-1/3)^2}{4} + \frac{(-1/3)^2}{4}} = 4.172; -11.83/4.172 = -2.84 < -2.101$, so companies B and C differ significantly at $P = 0.05$.

4.
(a) transformations can help to attain normality and/or constant variance of error distribution (but can’t help with independence); e.g. log transformation useful if residual vs fitted plot shows variance increasing with the mean.

(b) (i) both are measures of spread, standard deviations; $s$ is a sample value, a statistic, which can be observed; $\sigma$ is a population value, a parameter, and is generally unknown.
(ii) a CI is a statement about a population parameter, such as a population mean; a PI is a statement about a future observation; a PI will be wider; a CI narrows as the sample size increases.
(iii) both are probabilities; $P$-value is the probability of observing the outcome or a more extreme one under $H_0$, and is a sample value; power is the probability of rejecting $H_0$ under $H_1$, and is a property of the test.

5.
(a) I = RBD, block = row of trees (across the page); II = none, treatments not randomised to trees (only to columns), treatments confounded with columns; III = RBD, block = column of trees (down the page).

(b) III, because wetness gradient will be confounded with blocks, not treatments; trees within blocks should be of similar wetness.

(c) Source DF
---
Blocks 4
Treatments 4
Residual 16
Total 24
6. (a) response: number of items sold, numerical, discrete; explanatory: price, display, both categorical, ordinal.

(b)

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>price</td>
<td>2</td>
<td>30778</td>
<td>15389</td>
<td>3172</td>
</tr>
<tr>
<td>display</td>
<td>2</td>
<td>17006</td>
<td>8503</td>
<td>1739</td>
</tr>
<tr>
<td>Interaction</td>
<td>4</td>
<td>5087</td>
<td>1272</td>
<td>260</td>
</tr>
<tr>
<td>Error</td>
<td>18</td>
<td>88</td>
<td>4.89</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>26</td>
<td>52959</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$F = 260$, so interaction highly significant; consider top left 2 × 2 subset: at reduced price, normal display results in a mean increase of 40 items, while at regular price, there is virtually no increase.

(c) $((153.33 - 154.74)^2 + (114.11 - 154.74)^2 + (196.78 - 154.74)^2) \times 9 = 30781 \approx 30778$. 

(d) fitted value = 189.67, so residual = 186 − 189.7 = −3.67.

(e) $121.33 - 120.00 \pm t_{18}(.975) \times \sqrt{1.89/1/3 + 1/3} = 1.33 \pm 2.101 \times 1.806 = (-2.46, 5.12)$. Large or normal display — these are not significantly different.

7. (a) for every additional pound of left leg strength, the ball is expected to stay in the air an extra 0.0135 seconds, provided the other variables remain constant.

(b) $x_2$ has smallest res SS; $F = \frac{(2.868-0.736)/(12-11)}{0.736/11} = 31.9$, so add $x_2$; $x_2, x_3$ has smallest res SS; $F = \frac{(0.736-0.369)/(11-10)}{0.369/10} = 9.95$, so add $x_3$; don’t add $x_1$ ($F = 0$); so select $x_2$ and $x_3$. $x_1$ is eliminated ($F = 0$); $x_2$ has smallest res SS, but $x_3$ can’t be eliminated, since it came in with forward selection, so model is the same.

(c) $x_1$ and $x_2$ may be highly correlated (especially as they are similar variables), so once one of them is in the model, the other doesn’t contribute much.
8. 
(a) \( \mu(i,x) = \mu + \alpha_i + \beta_i x \) \( \mu(i,x) = \mu + \alpha_i + \beta x \) (or equivalent formulations); second model more appropriate since interaction not significant \( P = 0.607 \).

(b) graph of SCC vs preSCC should have parallel lines, Antibact line tangibly below Untreated line.

(c) \( \mu(1,x) = 172.6 + 94.43 + 0.5916x = 267.03 + 0.5916x \); predicted value = \( 267.03 + 0.5916 \times 808 = 745.0 \), so residual = \( 931 - 745.0 = 186.0 \); 186.0/62 = 3.0, which is large enough to be an outlier.

(d) Antibact significantly reduces the SCC of milk; SCC the day before has a significant effect on SCC.

9. 
(a) \( Y_{ijk} = \mu + \alpha_i + B_j + (\alpha B)_{ij} + E_{ijk} \), where \( \alpha_i \) = (fixed) effect of font \( i \), \( B_j \) = (random) effect of student \( j \), \( (\alpha B)_{ij} \) = interaction effect of font \( i \) and student \( j \), \( E_{ijk} \) = random error.

(b) \( \sigma^2 + 2\sigma^2_{\alpha B} = 27.43 \), \( \sigma^2 + 2\sigma^2_{\alpha B} + 6\sigma^2_B = 148.08 \), so \( \sigma^2_B = (148.08 - 27.43)/6 = 20.11 \).

(c) \( F = 148.08/27.43 = 5.40 > F_{4,8}(0.95) = 3.84 \), so variance component is significant.

(d) Could be random if experimenters not interested in particular fonts, and only concerned about variation between fonts.