This is a open-book, open-notes, computer-assisted examination. Please write up your results in report form using Word or L\textsc{a}T\textsc{e}X, and turn in a hard copy to the proctor or email a .pdf or .doc file to hatfield@umn.edu by 3:55 pm today. That is, you will have 85 minutes (2:30–3:55 pm) to complete this data analysis and write up its results. Good luck.

<table>
<thead>
<tr>
<th>week ($X_j$)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>game 1 ($Y_{1j}$)</td>
<td>85</td>
<td>78</td>
<td>64</td>
<td>94</td>
<td>87</td>
<td>110</td>
<td>122</td>
<td>100</td>
<td>110</td>
<td>123</td>
<td>105</td>
<td>112</td>
<td>103</td>
<td>140</td>
</tr>
<tr>
<td>game 2 ($Y_{2j}$)</td>
<td>136</td>
<td>142</td>
<td>87</td>
<td>129</td>
<td>113</td>
<td>65</td>
<td>110</td>
<td>103</td>
<td>92</td>
<td>133</td>
<td>118</td>
<td>94</td>
<td>132</td>
<td>125</td>
</tr>
<tr>
<td>game 3 ($Y_{3j}$)</td>
<td>92</td>
<td>127</td>
<td>136</td>
<td>149</td>
<td>119</td>
<td>148</td>
<td>120</td>
<td>159</td>
<td>127</td>
<td>121</td>
<td>117</td>
<td>153</td>
<td>118</td>
<td>108</td>
</tr>
</tbody>
</table>

1. Katie bowled $G = 3$ games each week for $W = 14$ weeks; the data are given in the table above and online at www.biostat.umn.edu/~brad/data/bowling_data.txt. Shown are the week, $X_j$, and the three weekly scores, $Y_{ij}$, for $i = 1, 2, 3$, and $j = 1, \ldots, 14$.

Katie’s first thought is to fit 3 separate (uncentered) simple linear regressions,

$$Y_{ij} = \alpha_i + \beta_i X_j + \epsilon_{ij}, \quad \epsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2),$$

for $i = 1, 2, 3$.

(a) Use \textsc{R} to plot the raw data. Do the time trends for the 3 games appear similar?  
[Hint: You may find it helpful to decorate your plot with traditional least squares lines using the \texttt{lsfit} and \texttt{abline} commands.]

(b) Fit the model in \textsc{WinBUGS}, using vague priors for all parameters. Find and compare the posterior density estimates and 95% credible intervals for week effect parameters $\beta_i$, $i = 1, 2, 3$. Do Katie’s scores significantly improve over time?

(c) Assuming your model is correct, obtain Bayesian point and 95% interval estimates for the 3 scores Katie would obtain were she to continue on for a 15th week. Compare and comment on any interesting features.

(d) Next, consider a simpler model which pools the data from all games, i.e.,

$$Y_{ij} = \alpha + \beta X_j + \epsilon_{ij}, \quad \epsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2).$$

Find and comment on the posterior of the overall time trend $\beta$. Also compare the effective model sizes $p_D$ and DIC scores for your two models. Which do you prefer, and why?

(e) Finally, to avoid the somewhat restrictive assumption that the scores are linear in $X$, try a more complex two-way analysis of variance (ANOVA) model,

$$Y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}, \quad \epsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2).$$

To ensure parameter identifiability, use vague but not flat priors for the row and column effects, e.g., $\alpha_i \overset{iid}{\sim} N(0, \sigma^2_\alpha)$ and $\beta_j \overset{iid}{\sim} N(0, \sigma^2_\beta)$ where $\sigma^2_\alpha = \sigma^2_\beta = 100^2$. Check convergence of your model and comment on any issues. Then once again compare $p_D$ and DIC for this ANOVA model with those from your two regression models. Is the additional ANOVA-style complexity worth it?
[Overall Hint: If you decide to go for simultaneous fitting of all three models, as illustrated in class using the “binary dugongs” example,

www.biostat.umn.edu/~brad/data/dugongsBin_BUGS.txt,

be careful to use different names for your data and parameters across models. That is, don’t just re-use names like alpha and sigma over and over again, as this will confuse WinBUGS!]