1. a) 95% CI for proportion of blue candle

\[ \text{Sample} \]
5 blue
39 not blue.

\[ \text{Expected proportion} = \frac{5}{44} = 0.1136 \quad \text{or} \quad 11.36\% \]

95% CI is \( \text{EV} \pm 2SE \)

\[ SE_{\text{proportion}} = \frac{SE_{\text{sum}}}{\# \text{draws}} = \frac{\sqrt{\# \text{draws} - \text{SD}_{\text{sum}}}}{\text{SD}_{\text{sum}}} \]

\[ = \frac{\sqrt{44 \times (1-0) \frac{5}{44} \times \frac{39}{44}}}{44} = 0.0478 \times 5\% \]

95% CI is 1.8% to 20.8%.

5) (i) Two white candles

\[ \text{p (first is white) } \times \text{p (second is white | first is white)} \]

\[ \frac{19}{44} \times \frac{18}{43} = \frac{342}{903} = 0.3808 \quad 0.1808 \]

(ii) Two of the same color

\[ \text{p (two white) } + \text{p (two red) } + \text{p (two blue) } + \text{p (two yellow)} \]

[These are mutually exclusive events]

\[ \frac{19}{44} \times \frac{18}{43} + \frac{11}{44} \times \frac{10}{43} + \frac{5}{44} \times \frac{4}{43} + \frac{9}{44} \times \frac{8}{43} \]

\[ = \frac{544}{1892} = 0.2875 \]
(iii) Two different coloured candles.

\[ P(1^{\text{st}} \text{ candle is white and } 2^{\text{nd}} \text{ is not white}) + P(1^{\text{st}} \text{ red and } 2^{\text{nd}} \text{ is not red}) + P(1^{\text{st}} \text{ blue and } 2^{\text{nd}} \text{ is not blue}) + P(1^{\text{st}} \text{ yellow and } 2^{\text{nd}} \text{ is not yellow}) = \]

\[ \frac{19}{44} \times \frac{(43-18)}{43} + \frac{11}{44} \times \frac{33}{43} + \frac{5}{44} \times \frac{37}{43} + \frac{9}{44} \times \frac{35}{43} = \]

\[ = \frac{1848}{1892} = 0.7125 \]

Note: \( P(\text{two different coloured candles}) = 1 - P(\text{two candles of same colour}) \)

\[ = 1 - 0.2875 \]
\[ = 0.7125. \]

c) No - proportion of each colour in population is 0.25

\( H_{1} \) - some colours are morelikely.

<table>
<thead>
<tr>
<th>Colour</th>
<th>Observed</th>
<th>Expected ( \frac{44 \times 16}{43} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>white</td>
<td>19</td>
<td>11</td>
</tr>
<tr>
<td>red</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>blue</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>yellow</td>
<td>9</td>
<td>11</td>
</tr>
</tbody>
</table>
\[ \chi^2 = \frac{(19-11)^2}{11} + \frac{(11-11)^2}{11} + \frac{(5-11)^2}{11} + \frac{(7-11)^2}{11} \]

\[ = 9.45 \]

Fully specified model \( \Rightarrow 4 - 1 = 3 \) dof.

From \( \chi^2 \) table, \( 9.45 \) is between 5\% and 1\%.

Reject the at the 5\% level.

The data support my older son's position.
\[ \frac{1}{26^7} = \frac{1}{8.03 \times 10^9} = 1.245 \times 10^{-10} \quad (1 \text{ in } 8 \text{ billion}). \]

b) \[
\begin{align*}
0.02228 \times 0.02753 \times 0.042782 \times 0.00772 \times \\
0.01974 \times 0.07507 \times 0.02753
\end{align*}
\]
\[= 5.3958 \times 10^{-12}. \quad (1 \text{ in } 1.85 \times 10^{11}). \]

c) \[
\begin{align*}
0.03779 \times 0.01487 \times 0.03511 \times 0.0069 \times \\
0.0160 \times 0.06264 \times 0.01487
\end{align*}
\]
\[= 2.929 \times 10^{-12} \quad (1 \text{ in } 4.93 \times 10^{11}). \]
\[= 1 \text{ in } 493 \text{ billion}. \]

d) \[26^7 = 8.03 \times 10^9\]

e) chance is \[\frac{\text{# that we're interested in}}{\text{total # of combinations}} = \frac{24029}{8.03 \times 10^9} = 3 \times 10^{-6} \]
\[= 1 \text{ in } 330,000. \]
f) No - the chance (even if quit (e)) is too small.
(3)

a) Controlled experiment - the participants were assigned randomly to treatment or control.

b) It was not performed blind - the participants knew whether they were asked to lie down or not.

**Double blind** - double knowledge. It is possible that the patients knew no, the staff knew whether they were asking the participant to lie down or not.

Also, however, the staff diagnosing pregnancy may not have known whether the participant was in the treatment or control group.

It is not an important consideration here, as

(i) diagnosing pregnancy is not subjective

(ii) you can’t “think yourself” pregnant (placebo effect).

c) No - provided that the randomization to immobilization/mobilization was performed the same at each location.

d) Immobilization increases pregnancy rates

e) H0: rates of pregnancy are the same for patients who were immobilized as for patients who were not.

H1: rates of pregnancy are higher for immobilized patients.
i) non-mobile: 199 couples ≤4 became pregnant

mobile: 192 couples ≥4 became pregnant.

2-sample z-test.

\[ z = \frac{\text{observed diff} - \text{expected diff}}{\text{SE of diff}} \]

non-mobile: \[ \frac{54}{199} \times 100\% = 27.14\% \text{ pregnant} \]

\[ \text{SE}_{%} = \frac{\overline{x}_{199} \times 100}{\sqrt{\text{draws}}} = \sqrt{\frac{54}{199} \times \frac{145}{199}} \times 100 = 3.15\% \]

mobile: \[ \frac{34}{192} \times 100\% = 17.71\% \text{ pregnant} \]

\[ \text{SE}_{%} = \sqrt{\frac{34}{192} \times \frac{158}{192}} \times 100 = 2.735\% \]

\[ \text{SE}_{\text{diff}} = \sqrt{3.15^2 + 2.735^2} = 4.19\% \]

\[ z = \frac{(54 - 34) - 0}{4.19} = 4.77 \]

can conclude at 7% significance level

(\text{or 0.1, or 0.01...})

g) because even though the populations are not independent, and not sampling with replacement, 2-sample z-test works for randomized controlled trials.
a) correlation bet. quality and year is positive

b) mean year = 1979.5

\[
\frac{2000}{2000 - 1979.5} = \frac{2.0582}{2.94} \quad \text{SD year about mean.}
\]

predicted # frosts = \[ 8.2 + (-0.57) \times 2.0582 \times 6.4 \]

\[ = 0.6917 \]

RMS error is \[ \sqrt{1 - r^2} \times \text{SD of frosts} \]

\[ = \sqrt{1 - 0.57^2} \times 6.4 \]

\[ = 5.2583 \]

c) fitness.

will not correspond to reality as can't have a negative number of frosts.
f) If the outlier was improved the correlation would increase as the remaining points are more tightly clustered.
9) To exclude an outlier, there needs to be some extra information (e.g., recognising a faulty recording) to justify ignoring it.

10) decreases. When averages are used (as here), some variability has been removed. Getting that variability back (as would decrease the magnitude of the correlation).

b) Predicted fit (cont.) 0.6917.

Thus \(-8.2 - 0.6917 \times \frac{-1.1732}{6.6} = 0.1732\) SD below mean

Predict \(86 + (-0.64) \times (-1.1732) \times 5.6 = 90.2047 \approx 90\)

as the quality in 2000