Concept questions

Concept question 1. Overnight polling
During the presidential election season, pollsters often do ‘overnight polls’ and report a ‘margin of error’ of about ±4%.

The number of people polled is in which of the following ranges?
(a) 0 – 50
(b) 50 – 100
(c) 100 – 500
(d) 300 – 600
(e) 600 – 1000

Board questions

Problem 1. Confidence intervals for a binomial
For a poll to find the proportion \( \theta \) of people supporting X we know that a \( (1-\alpha) \) confidence interval for \( \theta \) is given by

\[ \bar{x} - \frac{z_{\alpha/2}}{\sqrt{n}}, \quad \bar{x} + \frac{z_{\alpha/2}}{\sqrt{n}} \]

(a) How many people would you have to poll to have a margin of error of 0.01 with 95% confidence? (You can do this in your head.)

(b) How many people would you have to poll to have a margin of error of 0.01 with 80% confidence. (You’ll want R or other calculator here.)

(c) If \( n = 900 \), compute the 95% and 80% confidence intervals for \( \theta \).

Problem 2. Pivoting: confidence intervals and non-rejection regions
This question gets at the relationship between confidence intervals and non-rejection regions.

Main point: For a sample with sample mean \( \bar{x} \), the confidence interval consists of all values \( \mu \) for which a NHST with null hypothesis mean = \( \mu \) would not reject on seeing \( \bar{x} \).

Assume we have independent data \( x_1, \ldots, x_n \sim N(\mu, \sigma^2) \), where \( \mu \) is unknown and \( \sigma \) is known.

(a) For null hypothesis \( \mu = \mu_0 \) give the two-sided non-rejection region for significance level \( \alpha \).

(b) Call the data average \( \bar{x} \). Give the \( 1-\alpha \) confidence interval for \( \mu \).

(c) Use the \( \bar{x}, \mu \)-plane below. Note the conveniently included guides.

(i) Plot the horizontal line segment at height \( \mu_0 \) showing the non-rejection region for \( H_0 : \mu = \mu_0 \) (significance level = \( \alpha \)).

(ii) Plot the horizontal line segment at other heights showing the non-rejection region for the corresponding \( \mu \).
(iii) Plot the vertical line segments showing the $1 - \alpha$ confidence intervals around $\overline{x}_1$ and $\overline{x}_2$.

(iv) Plot the vertical line segment at other values of $\overline{x}$ showing the corresponding confidence interval.

Understand how the main point connects with your graph.

Problem 3. Exact binomial confidence interval
This was not used in class, but it is a nice problem, so we included it here.
Use this table of binomial$(8, \theta)$ probabilities to:

1. find the (two-sided) rejection region with significance level 0.10 for each value of $\theta$.
2. Given $x = 7$, find the 90% confidence interval for $\theta$.
3. Repeat for $x = 4$.

<table>
<thead>
<tr>
<th>$\theta \setminus x$</th>
<th>0</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>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.430</td>
<td>0.383</td>
<td>0.149</td>
<td>0.033</td>
<td>0.005</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>0.3</td>
<td>0.058</td>
<td>0.198</td>
<td>0.296</td>
<td>0.254</td>
<td>0.136</td>
<td>0.047</td>
<td>0.010</td>
<td>0.001</td>
<td>0.000</td>
</tr>
<tr>
<td>0.5</td>
<td>0.004</td>
<td>0.031</td>
<td>0.109</td>
<td>0.219</td>
<td>0.273</td>
<td>0.219</td>
<td>0.109</td>
<td>0.031</td>
<td>0.004</td>
</tr>
<tr>
<td>0.7</td>
<td>0.000</td>
<td>0.001</td>
<td>0.010</td>
<td>0.047</td>
<td>0.136</td>
<td>0.254</td>
<td>0.296</td>
<td>0.198</td>
<td>0.058</td>
</tr>
<tr>
<td>0.9</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.005</td>
<td>0.033</td>
<td>0.149</td>
<td>0.383</td>
</tr>
</tbody>
</table>

Problem 4. Pivoting: Chi square confidence intervals for variance
This was not used in class, but it is a nice problem, so we included it here.
Assume we have independent data $x_1, \ldots, x_n \sim N(\mu, \sigma^2_{\text{true}})$, where $\sigma_{\text{true}}$ is unknown and our parameter of interest.

Let $s^2$ be the sample variance. We know that $\frac{(n-1)s^2}{\sigma^2_{\text{true}}} \sim \chi^2(n - 1)$. Thus,

$$P(c_{1-\alpha/2} \leq \frac{(n-1)s^2}{\sigma^2_{\text{true}}} \leq c_{\alpha/2} | \sigma_{\text{true}} = \sigma) = 1 - \alpha.$$ 

Here, $c_{\alpha/2}$ is the right critical point for the $\chi^2(n - 1)$ distribution.
Using this, for a two-sided significance test with $H_0 : \sigma_{\text{true}} = \sigma$, the non-rejection region for $s^2$ at significance level $\alpha$ is

$$\frac{c_{1-\alpha/2} \sigma^2}{n-1} \leq s^2 \leq \frac{c_{\alpha/2} \sigma^2}{n-1}$$

Pivoting, we get the $1-\alpha$ confidence interval for $\sigma^2$ produced by the data is

$$\frac{(n-1)s^2}{c_{\alpha/2}} \leq \sigma^2 \leq \frac{(n-1)s^2}{c_{1-\alpha/2}}$$

Display this graphically on the $\sigma^2$-$s^2$ axes shown.