Estimation and Confidence Intervals

Math 363 - April 21, 2011

1. The independent random variables $X_1, X_2,\ldots, X_n$ have common cumulative probability distribution function

$$F_X(x|\alpha, \beta) = \begin{cases} 0 & \text{if } x < 0, \\ (x/\beta)^\alpha & \text{if } 0 \leq x < \beta, \\ 1 & \text{if } x \geq \beta. \end{cases}$$

(a) Find the density of these random variables.
(b) The maximum likelihood estimator $\hat{\beta}$ for $\beta$ is the maximum value in the data. Find the maximum likelihood estimators for $\alpha$.
(c) The length (in millimeters) of cuckoos’ eggs found in hedge sparrow nests can be modeled with this distribution. Give the maximum likelihood estimates for the data

$$22.0 \ 23.9 \ 20.9 \ 23.8 \ 25.0 \ 24.0 \ 21.7 \ 23.8 \ 22.8 \ 23.1 \ 23.1 \ 23.5 \ 23.0 \ 23.0$$

2. Let $X_1, X_2,\ldots, X_n$ be a random sample with probability density function

$$f_X(x|\theta) = \theta x^{\theta-1}, \quad 0 \leq x \leq 1, \quad \theta > 0.$$ 

(This is a $Beta(\theta, 1)$ distribution.)

(a) Find the maximum likelihood estimator for $\theta$.
(b) Find the method of moments estimator of $\theta$.
(c) Use simulation to estimate the standard deviation of these estimators for $\theta = 1/2$ and $\theta = 2$.

3. The focal length of a lens in your lab has been mislabeled. You decide to make repeated measurements of the distance to the image and the object in order to estimate the focal length. Here $s_1$ is the distance from the lens to the object and $s_2$ is the distance from the lens to the real image of the object. Here are the measurements.

```r
> s1<-c(9.63, 9.925, 10.67, 10.30, 9.73, 9.60, 10.49, 9.43, 9.69, 10.20, 9.74, 9.44, 9.33, 10.60, 
+ 10.01, 9.78, 10.39, 10.08, 10.04, 9.55, 10.52, 10.24, 9.47, 9.86, 9.68)
> s2<-c(9.97, 9.90, 9.95, 9.94, 10.07, 10.14, 10.07, 10.16, 9.88, 10.06, 10.17, 10.13)
```

(a) Give a 98% confidence interval for $s_1$ and for $s_2$.
(b) The focal length $f$ is determined by using the thin lens formula,

$$\frac{1}{s_1} + \frac{1}{s_2} = \frac{1}{f}.$$ 

Give an estimate $\hat{f}$ based on these measurements and the thin lens formula.
(c) Use the delta method to give the standard deviation of \( \hat{f} \).
(d) Use this to devise a 98% confidence interval for \( f \).

4. A company with a fleet of 150 cars found that the emission system of 7 out of the 22 cars tested failed to meet pollution guidelines.

(a) Write a hypothesis to test if more than 20% of the entire fleet might be out of compliance.
(b) Test the hypothesis based on the binomial distribution and report a \( p \)-value.
(c) Is the test significant at the 10%, 5%, 1% level?

5. National data in the 1960s showed that about 44% of the adult population had never smoked.

(a) State a null and alternative hypothesis to test that the fraction of the 1995 population of adults that had never smoked had increased.
(b) A national random sample of 891 adults were interviewed and 463 stated that they had never smoked. Perform a \( z \)-test of the hypothesis and give an appropriate \( p \)-value.
(c) Create a 96% confidence interval for the proportion of adults who had never been smokers.
(d) Give the value of the power function \( \pi(p) \) for \( p = 0.46, 0.48, 0.50, 0.52 \) with the choice of \( \alpha = 0.04 \) and a “greater than” alternative hypothesis.
(e) Compute the power function for these values if we increase the sample to 1600. Explain why these values increased.

6. The body temperature in degrees Fahrenheit of 52 randomly chosen healthy adults is measured with the following summary of the data:

\[ n = 52, \quad \bar{x} = 98.2846, \quad s = 0.6824. \]

(a) Are the necessary conditions for constructing a valid \( t \)-interval satisfied? Explain.
(b) Find a 98% confidence interval for the mean body temperature and explain its meaning.
(c) Give a two-side hypothesis for a mean body temperature of 98.6°F Fahrenheit and use the information above to evaluate a test with significance level \( \alpha = 0.02 \).
(d) Find the power of the test at the parameter value \( \mu = 98.2 \) and indicate this value using the cutoff value for the test and drawing the sample distribution for the null and alternative hypothesis.

7. Drivers of cars calling for regular gas sometimes premium in the hopes that it will improve gas mileage. Here a rental car company takes 10 randomly chosen cars in its fleet and runs a tank of gas according to a coin toss, runs a tank of gas of each type.

<table>
<thead>
<tr>
<th>Car #</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>
</tr>
</thead>
<tbody>
<tr>
<td>Regular</td>
<td>16</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>22</td>
<td>27</td>
<td>25</td>
<td>27</td>
<td>28</td>
</tr>
<tr>
<td>Premium</td>
<td>19</td>
<td>22</td>
<td>24</td>
<td>24</td>
<td>25</td>
<td>25</td>
<td>26</td>
<td>26</td>
<td>28</td>
<td>32</td>
</tr>
</tbody>
</table>

(a) Write an appropriate hypothesis test for this situation and state the testing procedure appropriate to this circumstance.
(b) Compute the necessary summary statistics for the test in part (a).
(c) Perform the \( t \)-test and report the \( p \)-value.
(d) Compare your result to that of a two sample \( t \)-test.