

Assignment #4 — Due Wednesday, February 18, 2009, by 5:00 P.M.

Turn in homework in lecture, discussion, or your TA's mailbox. Indicate the discussion section in which you expect to attend to pick up this assignment on the assignment.

311: Monday 1:20–2:10**312:** Monday 12:05–12:55

- In the spirit of Valentine's Day, consider the following problem. A certain type of flower has three possible colors: red, pink, and white. Suppose that a sample of 400 flowers contains 35 red, 156 pink, and 209 white flowers.
 - The Hardy-Weinberg genetic model predicts that the relative proportions will be θ^2 , $2\theta(1 - \theta)$, and $(1 - \theta)^2$, respectively, where $\theta \in \Omega = (0, 1)$. Find a formula for the maximum likelihood estimate of θ as a function of the counts $s = (r, p, w)$. Find a numerical estimate of θ for this data. Plot the log-likelihood versus θ .

For making the plot, you may use the function `likelihoodPlot()` contained in the most recent version of the file `stat310.R`. This function requires a function you must write to compute the likelihood for this model. Please see the functions `locationNormal.logl()` and `poisson.logl()` in `stat310.R` as examples that you can copy and modify.
 - An alternative model predicts that the proportion of red, pink, and white flowers will have proportions θ_1 , θ_2 , and $1 - \theta_1 - \theta_2$ where $\theta = (\theta_1, \theta_2) \in \Omega = \{(\theta_1, \theta_2) : 0 \leq \theta_1, \theta_2, 1 - \theta_1 - \theta_2 \leq 1\}$. Sketch the set $\Omega \subset \mathbb{R}^2$. Find the MLE for θ . (You must solve a system of two equations.)
- Measurements in centimeters are assumed to be an i.i.d. normal sample. The sample ($n = 10$) is 4.7, 5.5, 4.4, 3.3, 4.6, 5.3, 5.2, 4.8, 5.7, and 5.3.
 - If we assume that the distribution for a single measurement is $N(\mu, \sigma_0^2)$ where $\sigma_0^2 = 0.5$, find a 90% confidence interval for μ .
 - If we assume that the distribution for a single measurement is $N(\mu, \sigma^2)$ where σ^2 is unknown, find a 90% confidence interval for μ .
- Suppose that $X_1, \dots, X_n \sim \text{Exponential}(\lambda)$.
 - Find formulas for the MLE and the standard error of the MLE.
 - Let $s = (3.96, 9.74, 3.50, 4.14, 1.73, 1.88, 18.49, 5.12)$ be modeled as the realization of an i.i.d. random sample from an Exponential distribution. Use the large sample MLE approximation $\hat{\lambda} \pm z\text{SE}(\hat{\lambda})$ where z is the quantile from a normal curve to construct an approximate 95% confidence interval for λ .
 - For the data in the previous part, graph the log likelihood versus lambda and mark the points on the curve at the boundaries of the confidence interval. Is the confidence interval of the form $\{\lambda \in \Omega : \log L(\lambda | s) \geq c\}$? Briefly explain.
- Use the `binomialSimulation()` in `stat310.R` to simulate 100 95% confidence intervals for θ from a `Binomial(20, 0.3)` model and plot the intervals. How many contain the true value of θ ?
- Repeat the previous problem for a sample of 10,000, but do not plot. How many intervals contain the true value of θ ?

Work to do, but not turn in.

- Read sections 6.3–6.5.
-