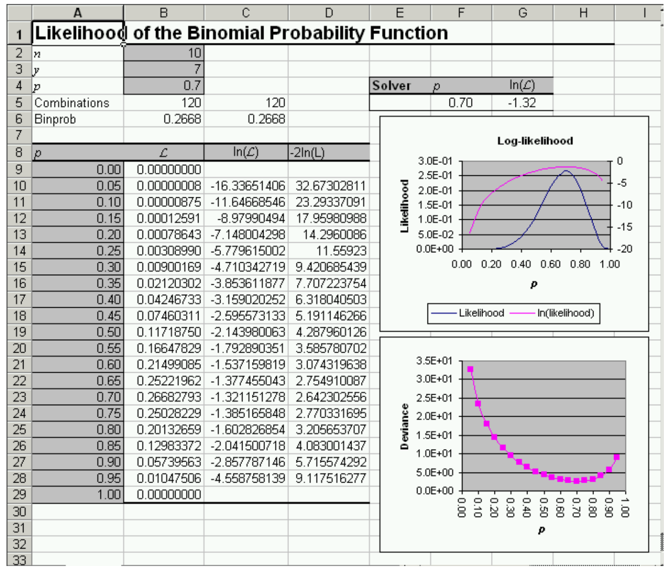
WILD 7250 - Analysis of Wildlife Populations

www.auburn.edu/~grandjb/wildpop

Part1—Likelihoods (50%)

1. Using the downloaded spreadsheet (lab_02.xls) as a template, create an Excel spreadsheet (including the graphs) similar to Figure 1 on page 1 of the Excel workbook. Note that the only cells containing values are the shaded ones. The remainders of the cells contain the results of calculations.

   Use the COMBIN function or the FACT to calculate the number of combinations and use BINOMDIST (set cumulative to FALSE in this function) and the explicit formula (Eq. 1) to calculate the probability of the result.



Likelihood of the Binomial Probability Function

	A	B	C	D
2	n	10		
3	y	7		
4	p	0.7		
5	Combinations	120	120	
6	Binprob	0.2668	0.2668	
7				
8	p	$\mathcal{L}$	ln($\mathcal{L}$)	-2ln(L)
9	0.00	0.00000000		
10	0.05	0.00000008	-16.33651406	32.67302811
11	0.10	0.00000875	-11.64668546	23.29337091
12	0.15	0.00012591	-8.97990494	17.95980988
13	0.20	0.00078643	-7.148004298	14.2960086
14	0.25	0.00308990	-5.779615002	11.55923
15	0.30	0.00900169	-4.710342719	9.420685439
16	0.35	0.02120302	-3.853611877	7.707223754
17	0.40	0.04246733	-3.159020252	6.318040503
18	0.45	0.07460311	-2.595573133	5.191146266
19	0.50	0.11718750	-2.143980063	4.287960126
20	0.55	0.16647829	-1.792890351	3.585780702
21	0.60	0.21499085	-1.537159819	3.074319638
22	0.65	0.25221962	-1.377455043	2.754910087
23	0.70	0.26682793	-1.321151278	2.642302556
24	0.75	0.25028229	-1.385165848	2.770331695
25	0.80	0.20132659	-1.602826854	3.205653707
26	0.85	0.12983372	-2.041500718	4.083001437
27	0.90	0.05739563	-2.857787146	5.715574292
28	0.95	0.01047506	-4.558758139	9.117516277
29	1.00	0.00000000		

Solver	p	ln($\mathcal{L}$)
	0.70	-1.32

Figure 1. Likelihood spreadsheet.

$$f(y \mid 10, 0.5) = \frac{n!}{y!(n-y)!} \times p^y \times (1-p)^{n-y}$$
$$= 120 \times 0.5^7 \times (1-.5)^{10-7} \qquad (1)$$
$$= 0.1172$$

The table in the spreadsheet is used to calculate the likelihoods using the formula below. Develop a formula to calculate the likelihood for values of p = 0, 0.05, 0.1... 1.0. (Eq. 2) Even though in this case it will not affect the outcome, include binomial coefficient in your calculations.

$$(p \mid n, y) = \binom{n}{y} p^y (1-p)^{n-y} \qquad (2)$$

Note that the log-likelihoods in the table can be calculated as either ln(_) or explicitly as:

$$\text{ln(\_) = ln(\_}(p \mid n, y)) = \ln \binom{n}{y} + y \cdot \ln(p) + (n-y) \cdot \ln(1-p)$$

The fourth column of the table -2 $\ln(\mathcal{L})$ is also known as the deviance. The deviance is distributed approximately as χ^2 and is a good measure of the fit of a model.

2. Examine the behavior of the functions

   Once you have the spreadsheet prepared, check it by comparing your calculations to those in the figure below. Then examine the following distributions and the comparable likelihoods:

$f(5\|10,0.5)$ (y = 5, n = 10, p = 0.5)	$f(2\|10,0.1)$
$f(25\|50,0.5)$	$f(20\|100,0.2)$
$f(1\|5,0.1)$	$f(40\|100,0.4)$
	$f(90\|100,0.9)$

   Note how the shape of the likelihood changes. The breadth of the peak is an indication of uncertainty in the model (i.e. it is related to variance) and is not affected by the value of $\hat{p}$, until $\hat{p}$ begins to approach 0 or 1.

3. Use the solver to determine the MLE

   NOTE: If the solver add-in does not appear in the Tools menu it must be installed from the original Excel installation disks.

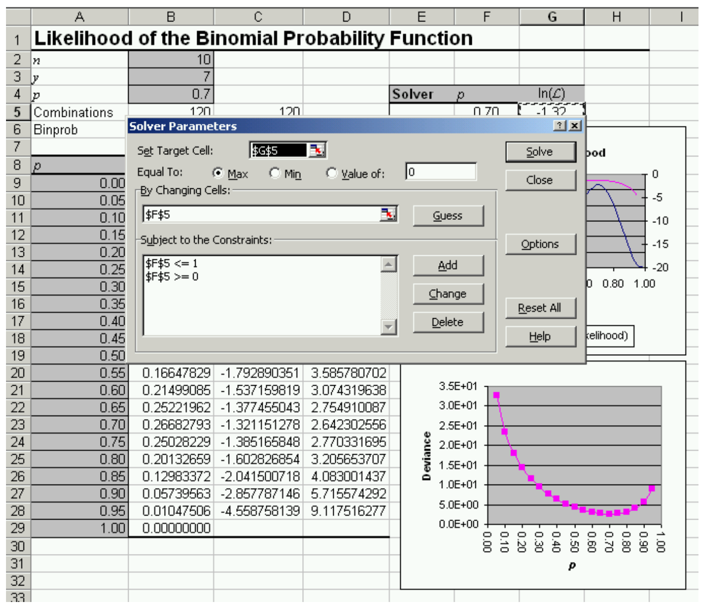


Figure 2. Using the solver.

   Using Solver:

   a. Start the solver from the Tools menu.

   b. Set the **target** cell to the cell containing the formula for the ln($\mathcal{L}$). On the Equal to line, click the **Max** radio button. In the space for by changing cells select the cell containing the **initial parameter estimate.** Add constraints so that the parameter estimate can not be $\leq$ 0 or $\geq$ 1. Click solve. Solver uses and iterative process of numerical estimation that is very similar to the ones employed in programs like SAS and MARK.

4. Create another worksheet similar to the one in the figure on the right with 20 observations from a binomial data set with 8 events and 12 nonevents. In the columns adjacent to the data calculate the probability of each observation and it corresponding natural logarithm. Row 24 in the figure are the product and the sum of the probabilities and their logs, respectively. Use the solver to determine the MLE for *p* (as in cell C3 in the figure).

	A	B	C	D	E
1	Likelihood for a Binomial Problem				
2					
3		***p***	0.27		
4		***n***	15		
5					
6					
7	**Obs**	***y***	**$\mathcal{L}$**	**ln($\mathcal{L}$)**	
8	1	1	0.2667	-1.3218	
9	2	0	0.7333	-0.3102	
10	3	0	0.7333	-0.3102	
11	4	1	0.2667	-1.3218	
12	5	0	0.7333	-0.3102	
13	6	0	0.7333	-0.3102	
14	7	0	0.7333	-0.3102	
15	8	0	0.7333	-0.3102	
16	9	1	0.2667	-1.3218	
17	10	0	0.7333	-0.3102	
18	11	0	0.7333	-0.3102	
19	12	0	0.7333	-0.3102	
20	13	0	0.7333	-0.3102	
21	14	1	0.2667	-1.3218	
22	15	0	0.7333	-0.3102	
23					
24		4	0.0002	-8.6987	
25					
26					

Part 2 – Information Theoretic Methods (50%)

The last three worksheets are used to calculate AIC and adjustments for small sample and overdispersion. These are the results of analysis of a data set we will used in future labs that examines the survival rate of radio-marked quail. In this study five models were cast including a null model (S(.)), and various combinations of group (g) and time (t) effects on survival.

1. The worksheet labeled AIC calculations contains the values for the effective sample size (n) from the study, and estimate of $\hat{c}$ and -2ln($\mathcal{L}$) and the number of estimated parameters (K). This is all of the information you will need as you fill in the remaining columns in the tables using the formulas from the lecture to calculate AIC, AICc, and QAICc the associated deltas (Δ), model weights, and model likelihoods.

   In the text box at the bottom of the worksheet, describe the behavior of AICc and QAICc and the associated deltas and model weights as sample size increases to 1000 and 10000 and as c-hat increases from 1 to 1.8 and 3, then briefly tell what this has to do with model uncertainty.

In the remaining worksheets, I dropped the null model (S(.)). Place your mouse pointer over cells marked by a red triangle in the upper right corner to read helpful comments.

2. The worksheet labeled multi-model inference contains the survival estimates and their respective standard errors for animals in group 1 during each time interval of the research project. You fill in the tables calculating the weighted average survival rates and unconditional estimates of the parameter standard errors

using the formulas provided in the lecture or from Anderson and Burnham (2000).

Note that the standard error is just the square root of the variance, and a good approximation of a 95% confidence limits is calculated as the estimate plus or minus twice the standard error.

In the text box at the bottom of the worksheet, describe how increasing model uncertainty would affect model averaged parameter estimates and unconditional estimates of the standard error. You may want to examine the effect by "tweaking" the model weights in column B.

3. Use the worksheet labeled parameter likelihood to calculate the probability that each parameter (group and time) belongs in the best approximating model. To do this, simply sum the model weights across the models that include that parameter. Note that this method is not recommended by Burnham and Anderson (2000) unless all models are equally represented in the model set.

Save the final worksheet as *yourlastname*_lab_02.xls and send it to grandjb@auburn.edu via e-mail before the next lab period.