Homework 3 Notes

First a couple of notes…

a. The average was the lowest of the three so far, though still about 93.5%. However, the problems I did see were generally more serious mistakes.
b. If you are losing several points on each of these assignments (especially if they are not just small things), you should really consider working with other students. Even better, work for a while on your own and then talk with other students. Some students are doing very well, and I know that some of these students would be willing to help those who are having trouble, especially after the student having trouble has put significant time in and gotten stuck. At the least, it is good to check your answers with someone to make sure you aren't way off base.
c. Again, if you have any questions, or you think I missed something, come to my office hours.

Problem 2.25 (10 points): Most of you did well on this problem, other than perhaps checking the second derivative to confirm that you found a maximum in part b. I took off a point for this just to make sure you remember to do it. I lost points for this on one of the tests.

Problem 2.28 (10 points): Again, most of you did well on this problem. I was ok if you accepted 2.27 as fact since that was not part of your assignment; however, for the test make sure you can find the answer without using 2.27a. Specifically, there were really three parts to problem.

1. You can factorize the likelihood to find that given beta, S=sum(y_i) is sufficient for alpha, answering the question as to why it turns out that the distribution of Y_k|S is free of the nuisance parameters alpha. This wasn't asked explicitly, so I didn't need to see this part.
2. To find that the likelihood is as given, you can use the result in 2.27a, but this result is just based on using the fact that f(Y|S)=f(Y,S)/F(S). The distribution of the numerator is just based on a binomial with p_i defined by the logistic distribution function and an indicator that the sum(y_i)=S. The distribution of the denominator is found in a similar way, except that we sum over all u where P(y_k1=u_1,y_k2=u_2,…,y_kmk=u_mk)…which is just the sum over all u of the multiplication of P(y_ki=u_i). Try to do this without using 2.27a if you used it on your homework.
3. As most of you identified, the likelihood is 1 when s_k=0 or m_k, and thus there is no information regarding beta.

Problem 2.33 (10 points): The key to this problem was to recognize that there are only k-1 uniquely defined parameters in the multinomial when there are k cells. If you try to find the information matrix based on k parameters, your answer will be wrong. Many of you who did this found that there is no covariance (the information matrix was 0 off the diagonal) among the cell totals. This certainly does not make sense intuitively. Even if you did find the information matrix based on k-1 parameters, some of you were unclear that the final matrix was only (k-1)x(k-1). Other problems were mostly algebraic. This information is found on page 65 in the book.

Problem 2.39 (10 points): If you had a problem on this one it was probably because you got a little confused about the likelihood in problem 2.12. The likelihood is in terms of the order statistics, not in terms of an independent exponential variable. So, you had to carry through both the order statistics y(i) and y(r) throughout the problem until you found the expected value using the hint given to you in the problem.

Some of you found the final answer that the information is r/(n*sigma^2). If you didn't simplify to this point, I did not take off since you weren't wrong, but you might be surprised to see that the final answer is so simple! However, if you take a couple minutes and write out the terms in the sums of your final answer, it should become clearly how to get r/(n*sigma^2). Specifically,

$$\sum_{i=1}^{r}\sum_{j=1}^{i}(n-j-1)^{-1}+(n-r)\sum_{i=1}^{r}(n-j-1)^{-1}=$$

$$\left(\frac{r}{n}+\frac{r-1}{n-1}+\frac{r-2}{n-2}+\cdots+\frac{1}{n-r+1}\right)+(n-r)*\left(\frac{1}{n}+\frac{1}{n-1}+\frac{1}{n-2}+\cdots+\frac{1}{n-r+1}\right)=$$

$1+1+1+\cdots+1=r$ (pair the first terms of each sequence).

This allows you to simplify your final answer to r/(n*sigma^2). There are a couple other ways to do this simplification using the sums, but this is the one I saw first.

Problem 2.43 (10 points): Most of you correctly identified that the derivative you want, dl/dbeta is just dl/dtheta*dtheta/dmu*dmu/dbeta and put the various parts together. There were a few ways to go about solving this using the chain rule but most of what I saw was correct as long as you took derivatives correctly and didn't get confused along the way.

Problem 2.44 (10 points): There were two main approaches here. First, if you took 1/n times the negative expected value of the second derivative of dl with respect to beta, you had to make sure to consider the derivative of all the parts involving beta (Di, mu_i, AND var(y_i)). After using the product rule and taking the derivative w.r.t. the first and third parts, you get a term with (y_i-mu_i) and then taking the expected value gives 0. Taking the derivative w.r.t. mu_i while leaving the other two constant gives you the desired information matrix after finding the expected value of this term.

The second approach involved taking 1/n times the sum of the expected value of the ith score equation multiplied by its transpose. Some of you consider the score equations in its entirety times its transpose which gives you a series of complicated sums (look at the definition of I-bar on page 62,2.39. Another common mistake was to write (y_i-mu_i) as a vector, which clearly isn't the case, though I didn't take off for this.