Theory of Inference: Homework 4

1. Here is a slightly technical question about hypothesis tests. Suppose that

$$Y_1, \dots, Y_n \overset{\text{iid}}{\sim} \text{Poisson}(\lambda) \qquad \lambda > 0.$$

The two competing hypotheses are

$$H_0 : \lambda = \lambda_0 \quad \text{versus} \quad H_1 : \lambda = \lambda_1$$

for specified values of λ_0 and λ_1.

(a) Derive the form of the likelihood ratio, $f_0(\boldsymbol{y})/f_1(\boldsymbol{y})$.

(b) Evaluate this ratio for the values $\lambda_0 = 0.50$, $\lambda_1 = 1.25$ and the dataset

$$\boldsymbol{y}^{\text{obs}} = \left\{1, 0, 2, 2, 0, 1, 1, 1, 1, 1, 3\right\}.$$

Thinking like a Bayesian, what does the dataset say about the two hypotheses?

(c) (Harder—maybe just think about this one over Easter.) Evaluate the Type 1 and Type 2 errors of the rejection region

$$\mathcal{R} := \left\{\boldsymbol{y} : \frac{f_0(\boldsymbol{y})}{f_1(\boldsymbol{y})} \le k\right\}$$

for $k \leftarrow 0.5$. Hint: you will find the Normal approximation to the Poisson very helpful for a pen-and-paper calculation (leave your results expressed in terms of the Normal distribution function), or else you can do the exact calculation in R, using the `ppois` function.

(d) (Another one to think about.) Produce a plot of the Type 1 error (x-axis) versus the Type 2 error (y-axis) for a range of values for k. This curve is known as the *operating characteristics curve* for this model and this pair of hypotheses. It is obvious that operating characteristics curve goes through $(0,1)$ and $(1,0)$, and it is fairly easy to prove that it is always convex (follows from the Neyman-Pearson Lemma).

2. Here is an exam-style revision question about hypothesis tests. Part 2d is unseen and a bit harder. Each part carries five marks.

   (a) Describe the general framework for Hypothesis Testing, and give illustrations of the types of hypotheses that might be tested.

   (b) Explain how a Bayesian might choose between two simple hypotheses.

   (c) State and prove the Neyman-Pearson Lemma.

   (d) Consider two simple hypotheses, H_0 and H_1, and the rejection region

   $$\mathcal{R} := \left\{ \boldsymbol{y} : \frac{f_0(\boldsymbol{y})}{f_1(\boldsymbol{y})} \leq k \right\}.$$

   Let α be the Type 1 error level of $\mathcal{R}$, and β be the Type 2 error level (these are both functions of k). Show that

   $$\frac{\Delta\alpha}{\Delta\beta} \approx -k,$$

   where Δ indicates the change that arises from a small change in k.

   (e) Explain briefly how we might examine the hypotheses

   $$H_0 : \theta = \theta_0 \quad \text{versus} \quad H_1 : \theta > \theta_0$$

   in the case of the model $\boldsymbol{Y} \sim \mathrm{p}(\cdot; \theta)$ with $\theta \geq 0$.

This homework is not due until the first Tue of next Term (21 Apr). In the meantime I will circulate another homework on p-values.

Jonathan Rougier
Mar 2015