ON THE MONOTONE LIKELIHOOD RATIO

OF A

"FAULTY INSPECTION" DISTRIBUTION

Leonard A. Stefanski
Department of Statistics
North Carolina State University
Raleigh, NC 27695

ABSTRACT

Let X denote the number of items *declared* defective in a sample of size n selected without replacemant from a batch of N items containing θ defectives, where testing is fallible. In this note it is shown that the density of X, $p(x, \theta)$, has a monotone likelihood ratio in x.

KEY WORDS: Faulty inspection distribution; Hypergeometric distribution; Measurement error model; Monotone likelihood ratio; Screening tests.

1. INTRODUCTION

A batch of N items contains an unknown number, θ, of defective items. A sample of n items is selected without replacement for testing. The procedure for testing items is fallible. With probability p_1 a defective item is declared defective, and with probability p_2 a nondefective item is declared nondefective. Both p_1 and p_2 are known, as are N and n. Let X represent the number of items declared defective among the n items tested. Then X has probability density function

$$p(x,\theta) = \sum_{\xi=0}^{\infty} \Pr(\xi \text{ defectives are chosen}) \times \left\{ \sum_{k=0}^{\infty} \Pr(k \text{ of } \xi \text{ defectives are declared defective}) \times \Pr(x-k \text{ of } n-\xi \text{ nondefectives are declared defective}) \right\}$$

$$= \sum_{\xi=0}^{\infty} \frac{\binom{\theta}{\xi}\binom{N-\theta}{n-\xi}}{\binom{N}{n}} \sum_{k=0}^{\infty} \binom{\xi}{k} p_1^k q_1^{\xi-k} \binom{n-\xi}{x-k} p_2^{n-\xi-x+k} q_2^{x-k}, \qquad (x = 0, 1, \ldots, n)$$

where $q_i = 1 - p_i$, $(i = 1, 2)$ and $\binom{a}{b} = 0$ when $b > a$ or $b < 0$.

In this note it is shown that $\{p(x,\theta) : \theta \in \Theta\}$ where $\Theta = \{0, 1, \ldots, N\}$ has a monotone likelihood ratio in x when $p_1 + p_2 > 1$. When $p_1 + p_2 = 2$, *i.e.*, the case of no misclassification, $p(x,\theta)$ is the standard hypergeometric distribution and provides a textbook example of a distribution with a monotone likelihood ratio, Lehmann (1986, p. 80).

The proof for general $1 < p_1 + p_2 < 2$ is more involved than that for the hypergeometric, possibly explaining why it is not readily found in the literature. Even the distribution $p(x,\theta)$ does not appear in classic works on discrete distributions, *e.g.*, Johnson and Kotz (1969); however, these authors have given a comprehensive description of $p(x,\theta)$ more recently (Kotz and Johnson, 1982).

The condition that $p_1 + p_2 > 1$ is familiar in the literature on screening tests, see Rogan and Gladen (1978), Gastwirth (1987) and Lew and Levy (1989). In that literature

p_1 and p_2 are known as the sensitivity and specificity respectively. A screening test with $p_1 + p_2 \leq 1$ is recognized as being useless.

It is well known that monotone likelihood ratio families admit "optimal" inference procedures (Cox and Hinkley, 1974, §§ 4.6 & 7.2; Lehmann, 1986, § 3.3; Casella and Berger, 1990, § 8.3.2). However, since X and θ are discrete, there are some technical difficulties exploiting the monotone likelihood ratio property in this case; see Lehmann (1986, p. 93) for a theoretical discussion of some of these issues, and Proctor (1990) for a practical resolution of these problems in an application involving $p(x,\theta)$. The latter paper motivated the present study of $p(x,\theta)$.

2. PROOF OF THE MONOTONE LIKELIHOOD RATIO PROPERTY

Above it was implied that proofs that $p(x,\theta)$ has a monotone likelihood ratio are not readily found in the literature. This is somewhat misleading since the proof given below can be considered a special case of Exercise 7 in Lehmann (1986, p. 530). However, the application of Lehmann's exercise is nontrivial, as will become evident in this section. Also, in Section 3 a connection to measurement error models is made. It is shown that the key step in the proof that $p(x,\theta)$ has a monotone likelihood ratio, is equivalent to demonstrating that X is an acceptable proxy for the unknown number of true defectives in the sample of size n.

Define

$$g(x,\xi,\theta) = \sum_{k=0} \binom{\xi}{k} \binom{n-\xi}{x-k} p_1^k q_1^{\xi-k} p_2^{n-\xi-x+k} q_2^{x-k},$$

and

$$h(\xi,\theta) = \frac{\binom{\theta}{\xi}\binom{N-\theta}{n-\xi}}{\binom{N}{n}}.$$

Note that $g(x,\xi,\theta)$ is constant in θ. It is defined this way in order to conform to the notation in Lehmann's exercise. With these definitions

$$p(x,\theta) = \sum_{\xi=0} h(\xi,\theta) g(x,\xi,\theta).$$

In accordance with Lehmann's terminology, a function $f(\cdot,\cdot)$ is *nondecreasing* provided

$$\frac{f(u',v)}{f(u,v)} \leq \frac{f(u',v')}{f(u,v')}$$

whenever $u \le u'$ and $v \le v'$. Note that $p(x,\theta)$ has a monotone likelihood ratio provided it is nondecreasing.

It follows from Lehmann's exercise that in order to establish that $p(x,\theta)$ is nondecreasing, it is sufficient to show that $g(x,\xi,\theta)h(\xi,\theta)$ is nondecreasing:

(i) in (x,θ) for ξ fixed;

(ii) in (θ,ξ) for x fixed;

and

(iii) in (x,ξ) for θ fixed.

To this end let $\theta' > \theta$, $\xi' > \xi$ and $x' > x$. Since

$$\frac{g(x,\xi,\theta')h(\xi,\theta')}{g(x,\xi,\theta)h(\xi,\theta)} = \frac{h(\xi,\theta')}{h(\xi,\theta)}$$

is independent of x, it is also nondecreasing in x thus proving (i).

Letting $\theta' = \theta + 1$ without loss of generality, the ratio

$$\frac{g(x,\xi,\theta')h(\xi,\theta')}{g(x,\xi,\theta)h(\xi,\theta)} = \frac{h(\xi,\theta')}{h(\xi,\theta)} = \frac{(\theta+1)(N-\theta-n+\xi)}{(\theta-\xi+1)(N-\theta)},$$

is nondecreasing in ξ, thus proving (ii).

It remains to show that

$$\frac{g(x,\xi',\theta)h(\xi',\theta)}{g(x,\xi,\theta)h(\xi,\theta)}$$

is nondecreasing in x for fixed θ. Again letting $\xi' = \xi + 1$ without loss of generality, and eliminating terms not dependent on x, shows that it is sufficient to establish that $g(x,\xi+1,\theta)/g(x,\xi,\theta)$ is nondecreasing in x for θ fixed.

Routine calculation shows that

$$\frac{g(x,\xi+1,\theta)}{g(x,\xi,\theta)} \propto \frac{\sum_{k=0} \binom{\xi+1}{k}\binom{n-\xi-1}{x-k}\left\{\frac{p_1p_2}{(1-p_1)(1-p_2)}\right\}^k}{\sum_{k=0} \binom{\xi}{k}\binom{n-\xi}{x-k}\left\{\frac{p_1p_2}{(1-p_1)(1-p_2)}\right\}^k},$$

where the constant of proportionality does not depend on x.

Define

$$d(x,\delta,\xi,n) = \sum_{k=0} \binom{\xi}{k}\binom{n-\xi}{x-k}\delta^k$$

and

$$q(x,\delta,\xi,n) = \frac{d(x,\delta,\xi+1,n)}{d(x,\delta,\xi,n)}.$$

Then

$$\frac{g(x,\xi+1,\theta)}{g(x,\xi,\theta)} \propto q(x,\delta,\xi,n)$$

for $\delta = p_1p_2/(1-p_1)(1-p_2)$. Note that $\delta > 1$ if and only if $p_1+p_2 > 1$.

It is now shown that when $\delta \geq 1$, $q(x,\delta,\xi,n)$ is nondecreasing in x using induction on $n = \xi, \xi+1, \ldots$.

When $n = \xi$,

$$q(x,\delta,\xi,n) = \frac{\sum_{k=0} \binom{\xi+1}{k}\binom{-1}{x-k}\delta^k}{\sum_{k=0} \binom{\xi}{k}\binom{0}{x-k}\delta^k} = \frac{0}{\binom{\xi}{x}\binom{0}{0}\delta^x} = 0,$$

which is nondecreasing in x.

Similar calculations show that when $n = \xi+1$,

$$q(x,\delta,\xi,n) = \frac{(\xi+1)\delta}{(1-\delta)x+\delta(\xi+1)},$$

which is nondecreasing in x when $\delta \geq 1$.

Now assuming that $q(x,\delta,\xi,\xi+r)$ is nondecreasing in x, it is proven that $q(x,\delta,\xi,\xi+r+1)$ is nondecreasing in x.

Using the identity $\binom{a+1}{b} = \binom{a}{b} + \binom{a}{b-1}$ it can be shown that

$$q(x,\delta,\xi,\xi+r+1) = \frac{d(x,\delta,\xi+1,\xi+r)+d(x-1,\delta,\xi+1,\xi+r)}{d(x,\delta,\xi,\xi+r)+d(x-1,\delta,\xi,\xi+r)}.$$

It follows that $q(x+1,\delta,\xi,\xi+r+1) - q(x,\delta,\xi,\xi+r+1) \geq 0$ provided

$$D_1 + D_2 + D_3 + D_4 \geq 0,$$

where

$$D_1 = d(x+1,\delta,\xi+1,\xi+r)d(x,\delta,\xi,\xi+r) - d(x+1,\delta,\xi,\xi+r)d(x,\delta,\xi+1,\xi+r);$$

$$D_2 = d(x+1,\delta,\xi+1,\xi+r)d(x-1,\delta,\xi,\xi+r) - d(x+1,\delta,\xi,\xi+r)d(x-1,\delta,\xi+1,\xi+r);$$

$$D_3 = d(x,\delta,\xi+1,\xi+r)d(x,\delta,\xi,\xi+r) - d(x,\delta,\xi,\xi+r)d(x,\delta,\xi+1,\xi+r);$$

$$D_4 = d(x,\delta,\xi+1,\xi+r)d(x-1,\delta,\xi,\xi+r) - d(x,\delta,\xi,\xi+r)d(x-1,\delta,\xi+1,\xi+r).$$

Note that $D_1 \geq 0$ when $q(x+1,\delta,\xi,\xi+r) - q(x,\delta,\xi,\xi+r) \geq 0$. Thus by the induction hypothesis, $D_1 \geq 0$. Using identical reasoning, the induction hypothesis is employed to show that $D_2 \geq 0$ and $D_4 \geq 0$. Since $D_3 \equiv 0$, it follows that $D_1 + D_2 + D_3 + D_4 \geq 0$, thus establishing that $q(x,\delta,\xi,\xi+r+1)$ is nondecreasing in x. This proves (iii) and completes the proof that $p(x,\theta)$ has a monotone likelihood ratio.

3. THE CONNECTION TO MEASUREMENT ERROR MODELS

If Ξ denotes the number of true defectives in the sample, then $X = \Xi + Z$, where Z can be regarded as the error made in measuring X. A minimal condition that X should satisfy in order to be an acceptable proxy for Ξ, is that $E\{\tau(\Xi) \mid X = x\}$ should be nondecreasing almost surely, whenever $\tau(\cdot)$ is nondecreasing. This follows from results in Efron (1965) and Lehmann (1966). The connection to measurement error models is spelled out in Hwang and Stefanski (1989).

In particular, if X is an acceptable proxy for Ξ, and Y is any random variable such that $E(Y \mid \Xi = \xi)$ is nondecreasing almost surely, then $E(Y \mid X = x)$ will also be nondecreasing almost surely. Thus substitution of X for Ξ as a predictor in a regression model will not alter monotonic relationships.

A sufficient condition for X to have the desired monotonicity-preserving property is that $p_{X|\Xi}(x' \mid \xi)/p_{X|\Xi}(x \mid \xi)$ is nondecreasing in ξ for all $x < x'$, (Efron, 1965; Lehmann, 1966; Hwang and Stefanski, 1989). It is readily seen that this condition is equivalent to Condition (iii) in Section 2. Thus X is an acceptable proxy for Ξ in the sense defined above.

Note that Condition (iii) involves only the conditional distribution of $X \mid \Xi$, and not the marginal distribution of Ξ.

REFERENCES

Casella, G. and Berger, R. L. (1990), *Statistical Inference*, Pacific Grove, California: Wadsworth & Brooks/Cole.

Cox, D. R. and Hinkley, D. V. (1974), *Theoretical Statistics*, New York: Chapman and Hall.

Efron, B. (1965), "Increasing properties of Polya frequency functions," *Annals of Mathematical Statistics*, 36, 272–279.

Gastwirth, J. L. (1987), "The statistical precision of medical screening procedures," *Statistical Science*, 2, 213–226.

Hwang, J. T. and Stefanski, L. A. (1989), "Monotonicity of regression functions in structural measurement error models," unpublished manuscript.

Johnson, N. L. and Kotz, S. (1969), *Distributions in Statistics - Discrete Distributions*, New York: John Wiley and Sons.

Kotz, S. and Johnson, N. L. (1982), "Errors in inspection and grading: distributional aspects of screening and hierarchal screening," *Communications in Statistics, Series A*, 11(18), 1997–2016.

Lehmann, E. L. (1966), "Some concepts of dependence," *Annals of Mathematical Statistics*, 37, 1137- 1153.

Lehmann, E. L. (1986), *Testing Statistical Hypotheses, Second Edition*, New York: John Wiley and Sons.

Lew, R. A. and Levy, P. S. (1989), "Estimation of prevalence on the basis of screening tests," *Statistics in Medicine*, 8, 1225–1230.

Proctor, C. H. (1990), "Enumerative inference with misclassification error," unpublished manuscript.

Rogan, W. J. and Gladen, B. (1978), "Estimating prevalence from the results of a screening test," *American Journal of Epidemiology*, 107, 71–76.