The Mode, Median, and Mean Inequality

RICHARD A. GROENEVELD AND GLEN MEEDEN*

Abstract

An elementary method of proof of the mode, median, and mean inequality is given for skewed, unimodal distributions of continuous random variables. A proof of the inequality for the gamma, F, and beta random variables is sketched.

KEY WORDS: Mode-median inequality; Median-mean inequality; Skewed distributions.

1. Introduction

We present here an elementary proof of the mode, median, and mean (expectation) inequality for unimodal, moderately asymmetric distributions of continuous random variables. The fact that these parameters lie in the order stated for positively skewed distributions is often assumed. However, proofs of these relationships do not appear in the mathematical statistics texts which we have examined. We shall briefly outline the general method of proof and leave the proofs for the gamma, F, and beta distributions as an exercise.

2. The Method of Proof for Positively Skewed Distributions

Let X be a nonnegative unimodal random variable with a continuous probability density function (pdf) $f(x)$ and cumulative distribution function (cdf) $F(x)$. Assume $M > 0$ is the mode of the random variable. Denote the median of the distribution by m and the expectation of X by μ. We rotate the curve of the pdf $f(x)$ from 0 to M about $x = M$. Let $g(x) = f(x)$, $0 \le x \le M$, and $g(x) = f(2M - x)$ on $M \le x \le 2M$. If $g(x) \le f(x)$ for $M \le x \le 2M$, with the strict inequality holding for at least one x, then $m > M$; since if $m \le M$, then

$$1 = \int_0^\infty f(x)\,dx > \int_0^{2M} g(x)\,dx \ge 1,$$

which is a contradiction. The latter inequality arises from the assumption that the area under $f(x) = g(x)$ on $[0, M]$ is at least one-half. This is shown geometrically in Figure 1. For the exercises that follow, it may prove algebraically simpler to use the substitution $x = y + M$ and show that

$$g(x)/f(x) = f(M - y)/f(M + y) = r(y) \le 1,$$

for $0 \le y \le M$, with strict inequality holding for at least one y.

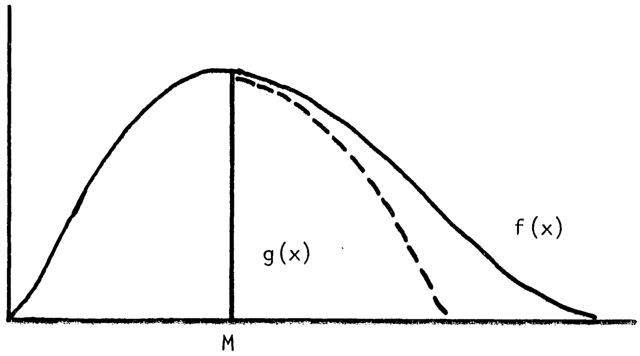

Figure 1

Once we find that $M < m$, we rotate the curve of $f(x)$ from 0 to m about the median. We define $h(x) = f(x)$, $0 \le x \le m$, and $h(x) = f(2m - x)$ for $m \le x \le 2m$. Clearly $h(x)$ is a density function of a random variable, say X_1. As $h(x)$ is symmetric about m, $E(X_1) = m$. Defining the cdf of X_1 to be $H(x)$ we see that $H(x) = F(x)$ for $0 \le x \le m$. If $f(x)$ is a density function for which there exists a unique $x_1 > m$ such that $h(x) > f(x)$ for $m < x < x_1$, $h(x_1) = f(x_1)$, and $h(x) < f(x)$ for $x > x_1$ (see Figure 2), $1 - F(x) \ge 1 - H(x)$ for all x, with strict inequality for $x > m$. For $x \le m$ and $x > x_1$ the inequality is obvious. For $m < x \le x_1$ we have $F(x) < H(x)$. Hence for the positive random variables X and X_1 we have

$$\mu = \int_0^\infty (1 - F(x))\,dx$$

$$> \int_0^\infty (1 - H(x))\,dx = E(X_1) = m,$$

i.e., the expectation exceeds the median.

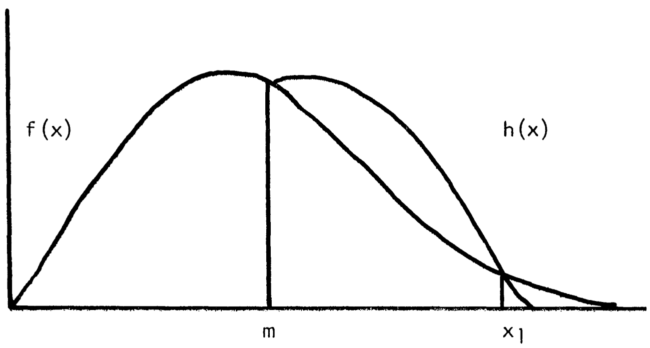

Figure 2

3. Application to Specific Random Variables

We consider three important families of positively skewed random variables with parameters restricted so that M and μ exist.

* Richard A. Groeneveld and Glen Meeden are with Department of Statistics, Iowa State University, Ames, IA 50011.

(a) The gamma distribution. The pdf for this class is

$$f(x) = x^{k-1}e^{-x/\beta}/\Gamma(k)\beta^k, \quad x > 0, k > 1, \beta > 0.$$

To show $M < m$, use

$$r(y) = f(M - y)/f(M + y)$$
$$= ((M - y)/(M + y))^{k-1}e^{2y/\beta}, \quad 0 \le y \le M. \quad (3.1)$$

Similarly, to show $m < \mu$, consider

$$h(x)/f(x) = (2m - x)^{k-1}$$
$$(e^{-(2m-x)/\beta})/x^{k-1}e^{-x/\beta}, \quad m \le x \le 2m.$$

Let $x = y + m$ and consider

$$t(y) = f(m - y)/f(m + y) \quad \text{on} \quad 0 \le y \le m. \quad (3.2)$$

The inequality for the chi-square distribution follows directly from the above where $\beta = 2$.

(b) The F distribution. The F distribution with (j, k) degrees of freedom has the pdf

$$j/k(\Gamma(j + k/2)/\Gamma(j/2)\Gamma(k/2))(jx/k)^{j/2-1}$$
$$(1 + jx/k)^{-(j+k/2)}, \quad x > 0, j > 2, k > 2,$$

and to show $m > M$ and $m < \mu$ consider $r(y)$ and $t(y)$ as defined in (3.1) and (3.2), respectively.

(c) The beta variable. Consider the beta random variable X with pdf

$$f(x) = (\Gamma(p + q)/\Gamma(p)\Gamma(q))x^{p-1}(1 - x)^{q-1},$$
$$0 \le x \le 1, q > p > 1,$$

and use arguments exactly analogous to those made for the gamma and F distributions.

4. Remark

While the method of proof and examples have addressed positively skewed distributions, it is clear by rotating the graphs in Figures 1 and 2 about the vertical axis that analogous results hold for negatively skewed distributions.

[*Received July 19, 1976. Revised February 23, 1977*]

References

Doodson, A. T. (1917), "Relation of the Mode, Median and Mean in Frequency Curves," *Biometrika*, 11, 425–9.

Haldane, J. B. G. (1942), "The Mode and Median of a Nearly Normal Distribution with Given Cumulants," *Biometrika*, 32, 294–9.

Lancaster, H. O. (1968), "The Median of χ^2," *Australian Journal of Statistics*, 10, 83.

Estimability and Linear Hypotheses

JUSTUS SEELY*

Abstract

A proposition is given which provides an easily justified reason as to why attention should be confined to estimable parametric vectors when formulating linear hypotheses. The possibility of justifying one's linear estimation effort on the estimable parametric functions via an identifiability condition is also mentioned.

KEY WORDS: Estimability; Identifiability; Linear hypothesis; Linear model; Testable hypothesis.

1. Introduction

The idea of an estimable parametric function in a linear model context was introduced in 1944 by Bose [1] who required that such a function possess a linear unbiased estimator. Since that time the topic has received considerable attention in the literature, and a discussion on estimability can be found in most textbooks which treat a linear model, $Y = X\beta + e$, where the mean vector $X\beta$ is overparameterized (i.e., where two different β vectors can lead to the same $X\beta$ vector). In most of these textbooks, however, there is little justification given, in testing a linear hypothesis of the form H_0: $H'\beta = \delta$ vs. H_1: $H'\beta \neq \delta$, as to why $H'\beta$ should be an estimable parametric vector, i.e., each component of $H'\beta$ is an estimable parametric function. (When $H'\beta$ is an estimable parametric vector, the linear hypothesis is often referred to as a testable hypothesis.) An exception is the book by Searle [6] where it is shown that unless $H'\beta$ is an estimable parametric vector, the numerator sum of squares in the F ratio is not well defined. The primary purpose of this article is to give an alternative justification for restricting attention to estimable parametric vectors in linear hypotheses. This alternative justification can be very simply stated as the requirement that the class of distributions under the null hypothesis be disjoint from the class of distributions under the alternative hypothesis.

In the process of establishing that a linear hypothesis should be formulated in terms of an estimable parametric vector, we give some additional known

* Justus Seely is with Department of Statistics, Oregon State University, Corvallis, OR 97331. The author wishes to thank Don Pierce for pointing out that the interpretation of the identifiability condition (a) is a natural way of introducing estimable parametric functions. The author also expresses his thanks to his colleague David Birkes, the referee, and the editor for helpful comments.