The Annals of Probability
1979, Vol. 7, No. 1, 96–108

SIGN CHANGES OF THE DIFFERENCE OF CONVEX FUNCTIONS AND THEIR APPLICATION TO LARGE DEVIATION RATES[1]

BY JAMES LYNCH

The Pennsylvania State University

The relationship of the large deviation rate, $\psi^*(a)$, of the mean of independent and identically distributed random variables to their cumulant generating function, $\psi(\lambda)$, is well known. This paper studies how the behavior of the sign changes of $\psi_1^*(a) - \psi_2^*(a)$ is related to that of $\psi_1(\lambda) - \psi_2(\lambda)$ for cumulant generating functions ψ_1 and ψ_2 with rates ψ_1^* and ψ_2^*, respectively. Use is made of the fact that the rate ψ^* is nothing more than the conjugate convex function of ψ. Results concerning the relationship of the behavior of the difference of convex functions to that of the difference of their conjugates are first proven and then applied to determine the relationship of the behavior of the sign changes of $\psi_1^* - \psi_2^*$ to that of $\psi_1 - \psi_2$. Results are also given relating this behavior to that of $F_1 - F_2$ and $f_1 - f_2$, where F_i and $f_i (i = 1 \text{ and } 2)$ are the distribution function and the density function, respectively, corresponding to ψ_i.

1. Introduction. Let $X_1, X_2, \cdots$ be independent identically distributed (i.i.d.) random variables with distribution F. Let $\phi(\lambda) = \int \exp(\lambda x) dF(x)$ and $\psi(\lambda) = \log \phi(\lambda)$, i.e., ϕ and ψ are, respectively, the moment generating function (m.g.f.) and the cumulant generating function (c.g.f.) of F. Let $\bar{X}_n = (X_1 + X_2 + \cdots + X_n)/n$. Then it follows from Chernoff's theorem (see Chernoff (1952) and Bahadur (1971)) that

$$n^{-1} \log P_F(\bar{X}_n \geqslant a) \to -\psi^*(a), \tag{1.1}$$

where

$$\psi^*(a) = \sup\{\lambda a - \psi(\lambda) : \lambda \geqslant 0\}. \tag{1.2}$$

From (1.2), it is clear that the *large deviation rate*, ψ^*, is directly related to the c.g.f. ψ. For two different distributions F_1 and F_2 with c.g.f.'s ψ_1 and ψ_2, respectively, and rates ψ_1^* and ψ_2^*, respectively, what is not clear, though, is how the behavior of $\psi_1^* - \psi_2^*$ is related to that of ψ_1-ψ_2. In this paper, we consider this problem. More specifically, we study how the behavior of the sign changes of $\psi_1^* - \psi_2^*$ is related to that of $\psi_1 - \psi_2$. To do this, we shall use the fact that ψ^* is nothing more than the conjugate convex function of ψ. In Section 3, certain basic properties of conjugate convex functions are discussed and results concerning the sign changes of their differences are proven. In Section 4, these results are used to determine the relationship of the behavior of the sign changes of $\psi_1^* - \psi_2^*$ to that of $\psi_1 - \psi_2$. Results are also given relating this behavior to that of $F_1 - F_2$ and $f_1 - f_2$, where f_i is the density of $F_i (i = 1 \text{ and } 2)$. The proofs of the results in

Received October 15, 1976; revised January 16, 1978.
[1]Research supported in part by a research initiation grant from The Pennsylvania State University.
AMS 1970 *subject classifications.* 26A51, 60R10, 62E10, 62E20.
Key words and phrases. Large deviations, conjugate convex functions, sign changes.

Section 4 rely quite heavily on the theory of total positivity. This, along with other basic concepts, is the topic of Section 2.

2. Sign changes and total positivity. Let $f(t)$ be a function defined on an interval I.

DEFINITION 2.1. The number of sign changes of $f(t)$ on I is

$$S_I(f) = \sup S[f(t_1), \cdots, f(t_m)],$$

where the supremum is extended over all sets $t_1 < t_2 < \cdots < t_m$ ($t_i \in I$), m is arbitrary but finite, and $S(x_1, \cdots, x_m)$ is the number of sign changes of the sequence $x_1, x_2, \cdots, x_m$, zero terms being discarded. When $S_I(f) < \infty$, $FS_I(f)$ will denote the final sign of f on I as the argument of f traverses I from left to right, zero terms being discarded. When the interval I is clearly understood, the subscripts will be suppressed.

DEFINITION 2.2. A function f is said to have a $+$ *to* $-$ ($-$ *to* $+$) *sign change* at x if there exists an $\epsilon > 0$ such that for each $\delta \in (0, \varepsilon)$,

$$f(x') \geqslant (\leqslant) 0 \quad \text{for} \quad x - \delta < x' < x$$

$$f(x') \leqslant (\geqslant) 0 \quad \text{for} \quad x < x' < x - \delta$$

and there are x_1 and x_2 with $x - \delta < x_1 < x < x_2 < x + \delta$ such that $f(x_2) < 0 < f(x_1)$ ($f(x_1) < 0 < f(x_2)$).

A powerful tool for the study of sign changes is the theory of total positivity. Below, we review some basic definitions and results from this theory which will be used in Section 4.

DEFINITION 2.3. Let X and Y be subsets of the real line. A function $K(x, y)$ on $X \times Y$ is said to be (strictly) *totally positive of order n* (TP_n) ((STP_n)) if

$$\begin{vmatrix} K(x_1, y_1) & \cdots & K(x_1, y_r) \\ \vdots & & \vdots \\ K(x_r, y_1) & \cdots & K(x_r, y_r) \end{vmatrix} \geqslant (>) 0 \tag{2.1}$$

for all $x_1 < x_2 < \cdots < x_r$ in X, $y_1 < y_2 < \cdots < y_r$ in Y, $r = 1, 2, \cdots, n$, where the term on the left of (2.1) denotes the determinant of the corresponding matrix.

A function which is (strictly) totally positive of all finite orders is said to be (strictly) *totally positive* (TP) ((STP)).

EXAMPLE 2.1. The function $K_1(x, y) = e^{xy}$ is STP in $x, y \in (-\infty, \infty)$ (see Karlin (1968), page 15). Thus, $K_2(r, t) = t^r$ is STP in $t \in (0, \infty)$ and $r \in (-\infty, \infty)$.

An important property of TP functions is their variation diminishing property which is stated in the next theorem. For a more general statement, a proof, and further discussion, see Karlin (1968), pages 21 and 233.

VARIATION DIMINISHING (VD) THEOREM. *Let $K(x, y)$ be a TP_r Borel measurable function on $X \times Y$ such that $\int_Y K(x, y)dy$ is finite for each $x \in X$. Let f be a bounded Borel measurable function on Y and let $g(x) = \int_Y K(x, y)f(y)dy$. Then $S(g) \leqslant S(f)$ provided $S(f) \leqslant r - 1$. Moreover, if*

$$S(g) = S(f) \leqslant r - 1,$$

then f and g exhibit the same sequence of signs when their respective arguments traverse the domain of definition from left to right.

3. Differences of convex functions and of their conjugates. Let $\psi(x)$ be an extended real value function defined on an interval I_0. Let $I = \{x \in I_0: \psi(x) < \infty\}$. The function ψ is said to be (strictly) convex on I_0 if I is an interval, ψ is (strictly) convex on I, and $\psi(x) = \infty$ for $x \in I_0 - I$.

We now define the conjugate convex function of a convex ψ.

DEFINITION. The function

$$\psi^*(y) = \sup\{xy - \psi(x): x \in I_0\}$$

is called the *conjugate convex function* or, simply, the *conjugate* of ψ.

It is clear from the definition of ψ^* that

$$\psi^*(y) = \sup\{xy - \psi(x): x \in I\}.$$

In the remainder of this section, the extended real valued convex functions (denoted by ψ's) will have domain $I_0 = [0, \infty)$ and satisfy:

(i) $\psi(0) = 0$;
and for $b = \sup\{x: x \in I\}$,
(ii) $\lim_{x \uparrow b} \psi(x) = \psi(b)$,
and
(iii) the derivative ψ' of ψ exists and is continuous on $[0, b)$ with $\psi'(0) = 0$ and $\lim_{x \uparrow b} \psi'(x) = \psi'(b)$, where we define $\psi'(b) = \infty$ if $b \not\in I$.

Here, and throughout the remainder of this paper, the derivative ψ' of a convex function ψ at the endpoints 0 and b ($b \in I$) will be defined by the appropriate one-sided derivative. Also, if $b < \infty$, let $\psi'(x) = \infty$ for $x > b$. Note that the derivative need not be finite at b even if $b \in I$.

REMARK. From (i) and (iii) ψ is nonnegative, increasing, and continuous on $[0, b)$. Also, since ψ is right continuous at 0 and satisfies (ii), ψ is said to be *closed* in the literature of convex functions. Closed convex functions and their conjugates have important properties which shall be exploited in the proofs of Theorems 3.1 and 3.2. See Roberts and Varberg (1973), Section 15; in particular, Theorems C and D. (Hereafter, we shall refer to these two theorems without reference to their source.)

Now let ψ_1 and ψ_2 be two extended real valued nonnegative strictly convex functions on $[0, \infty)$. For $i = 1$ and 2, let

$$I_i = \{x \in [0, \infty): \psi_i(x) < \infty\} \quad \text{and} \quad b_i = \sup\{x \in I_i\}.$$

Since $\psi_1(x)$ and $\psi_2(x)$ may both be ∞ at a point x, we adopt the convention that $\infty - \infty = 0$ for the purpose of sign changes.

The following two theorems indicate the relationship of $S_I(\psi_1 - \psi_2)$ and $S_{I'}(\psi_1^* - \psi_2^*)$ for certain types of intervals I and I'.

THEOREM 3.1. *Let ψ_1 and ψ_2 be two extended real valued strictly convex functions on* $[0, \infty)$ *which satisfy* (i), (ii) *and* (iii). *If there exists a positive $x^* \in I_1 \cap I_2$ such that $\psi_1'(x^*) = y^* = \psi_2'(x^*) < \infty$ and*

(iv) *$\psi_1 - \psi_2$ is not identically zero on any open subinterval of* $[0, x^*]$,

then

$$S_{[0, x^*]}(\psi_1 - \psi_2) = S_{[0, y^*]}(\psi_1^* - \psi_2^*). \tag{3.1}$$

Furthermore, when $S_{[0, x^*]}(\psi_1 - \psi_2) < \infty$, the values of $\psi_1^* - \psi_2^*$ and $\psi_2 - \psi_1$ exhibit the same sequence of signs (excluding zeroes) when their respective arguments traverse the domain of definition from left to right.

PROOF. If $S(\psi_1 - \psi_2) = 0$, then either $\psi_1(x) \geqslant \psi_2(x)$ for all $x \in [0, x^*]$ or $\psi_1(x) \leqslant \psi_2(x)$ for all $x \in [0, x^*]$. Without loss of generality, assume that $\psi_1(x) \leqslant \psi_2(x)$ for all $x \in [0, x^*]$. Then, for all $x \in [0, x^*]$ and $y \geqslant 0$,

$$xy - \psi_1(x) \geqslant xy - \psi_2(x).$$

Thus $\psi_1^*(y) \geqslant \psi_2^*(y)$ for $y \in [0, y^*]$ since the function $f_i(x) = xy - \psi_i(x)$ ($i = 1, 2$) on $[0, \infty)$ achieves its maximum on $[0, x^*]$ for each $y \in [0, y^*]$. Hence $S(\psi_1^* - \psi_2^*) = 0$ and $\psi_1^* - \psi_2^*$ has the same sign as $\psi_2 - \psi_1$.

Now we shall consider the case when $S(\psi_1 - \psi_2) \geqslant n$, where n is a positive integer.

Since $\psi_1 - \psi_2$ is not identically zero on any subinterval of $[0, x^*]$, we can choose n points $\{x_i\}_{i=1}^n$ with $0 = x_0 < x_1 < \cdots < x_n < x_{n+1} = x^*$ at which a sign change occurs and we can choose these points so that the sign change at x_i is opposite that at x_{i+1} for $i = 1, 2, \cdots, n-1$. Without loss of generality we assume that the sign change at x_1 is from $+$ to $-$.

Since $\psi_1(0) - \psi_2(0) = 0 = \psi_1(x_1) - \psi_2(x_1)$ and $\psi_1 - \psi_2$ has a $+$ to $-$ sign change at x_1, it follows from the continuity of $\psi_1 - \psi_2$ that $\psi_1 - \psi_2$ has a positive maximum on $[0, x_1]$ at a $w \in (0, x_1)$. So $\psi_1(w) > \psi_2(w)$ and $\psi_1'(w) = \psi_2'(w)$ since the maximum occurs on the interior of $[0, x_1]$. Let $c_1 = \psi_1'(w)$. Then by Theorem D,

$$\begin{aligned} \psi_1^*(c_1) - \psi_2^*(c_1) &= c_1 w - \psi_1(w) - (c_1 w - \psi_2(w)) \\ &= \psi_2(w) - \psi_1(w) < 0. \end{aligned} \tag{3.2}$$

Furthermore, since $\psi_1 - \psi_2$ has a $+$ to $-$ sign change at x_1 and $\psi_1(x_1) - \psi_2(x_1) = 0$, it follows that $\psi_1 - \psi_2$ has a negative minimum on $[x_1, x_2]$ at a $z \in (x_1, x_2]$. If $z \in (x_1, x_2)$, then $\psi_1(z) < \psi_2(z)$ and $\psi_1'(z) = \psi_2'(z)$ since the minimum occurs on the interior of $[x_1, x_2]$. If $z = x_2$, then x_2 is not a point of sign change since $\psi_1(z) - \psi_2(z) < 0$. Thus $x_2 = x^*$ and so $\psi_1'(z) = \psi_1'(z)$ by hypothesis. Let $c_2 = \psi_1'(z)$. Then for either of the above cases, i.e., $z \in (x_1, x_2)$ or $z = x_2$, an identity similar to (3.2) will show that $\psi_1^*(c_2) - \psi_2^*(c_2) > 0$.

If $n \geqslant 2$, we may repeat the above argument to show that there is a $z \in (x_{i-1}, x_i]$ $(i = 3, \cdots, n+1)$ such that $\psi_1'(z) = c_i = \psi_2'(z)$ and $\psi_1(z) > (<)\psi_2(z)$ if i is odd (even). So,

$$\psi_1^*(c_i) - \psi_2^*(c_i) < (>)0 \qquad \text{if} \quad i \quad \text{is odd (even)}.$$

From the monotonocity of ψ_1', $0 < c_1 < \cdots < c_{n+1}$. Thus, it follows that $S(\psi_1^* - \psi_2^*) \geqslant n$, and so,

$$S(\psi_1 - \psi_2) \leqslant S(\psi_1^* - \psi_2^*). \tag{3.3}$$

Now, since ψ_1 and ψ_2 satisfy (ii), it follows from Theorem D that ψ_i^* is a convex function which is strictly convex on $[0, y^*]$ with $\psi_i^{*\prime}(y^*) = x^*$, satisfies (i), (ii) and (iii), and $\psi_i^{**} = \psi_i$ for $i = 1$ and 2. Furthermore, it also follows that $\psi_1^* - \psi_2^*$ satisfies (iv) on $[0, y^*]$. Hence, from the above proof

$$S(\psi_1^* - \psi_2^*) \leqslant S(\psi_1 - \psi_2). \tag{3.4}$$

So, from (3.3) and (3.4), (3.1) holds.

It is clear from the above proof that when $S_{[0, x^*]}(\psi_1 - \psi_2) < \infty$, $\psi_1^* - \psi_2^*$ exhibits the same sequence of signs as $\psi_2 - \psi_1$. This completes the proof of Theorem 3.1.

For Theorem 3.2 we shall need the following notation. Let $b = \min\{b_1, b_2\}$ and let $W = \sup\{x \leqslant b\colon \psi_1'(x) = \psi_2'(x) < \infty\}$. Note that by (iii) W is well defined (though it may equal ∞) since $\psi_1'(0) = 0 = \psi_2'(0)$ and that $\psi_1'(W) = \psi_2'(W)$. Let $y^* = \psi_1'(W) = \psi_2'(W)$ and let $S = S_{[0, \infty)}(\psi_1 - \psi_2)$.

THEOREM 3.2. *Let ψ_1 and ψ_2 be two extended real valued strictly convex functions on* $[0, \infty)$ *which satisfy* (i), (ii), (iii) *and*

(iv′) $\psi_1 - \psi_2$ *is not identically zero on any open subinterval of* $[0, b)$. *Then, if* $2 \leqslant S < \infty$ *and W is less than or equal to the value at which the* $(S-1)$st *sign change occurs, then* $S_{[0, \infty)}(\psi_1^* - \psi_2^*)$ *is either equal to* $S-2$ *or* S; *and,* $\psi_1^* - \psi_2^*$ *and* $\psi_2 - \psi_1$ *exhibit the same first* $S - 1$ *signs (excluding zeroes) when their respective arguments traverse* $[0, \infty)$ *from left to right. In all other cases,* $S_{[0, \infty)}(\psi_1^* - \psi_2^*) = S$, *and if* $S < \infty$, $\psi_1^* - \psi_2^*$ *and* $\psi_2 - \psi_1$ *exhibit the same sequence of signs (excluding zeroes) when their respective arguments traverse* $[0, \infty)$ *from left to right.*

PROOF. If $b = 0$, the conclusion is immediate. So assume that $b > 0$.

We consider various cases.

CASE 1. $S = 0$. Proof is the same as the one given for this case in the proof of Theorem 3.1.

CASE 2. $S = \infty$. Since $S = \infty$, for each positive integer n, there is an $x_n > 0$ such that $S_{[0, x_n]}(\psi_1 - \psi_2) \geqslant n$ and $\psi_i(x_n) < \infty$ for $i = 1$ and 2. Hence, there exists an $x' \in (0, x_n]$ such that $\psi_1'(x') = y' = \psi_2'(x') < \infty$ and $S_{[0, x']}(\psi_1 - \psi_2) \geqslant n - 1$. Thus it follows from Theorem 3.1 that $S_{[0, y']}(\psi_1^* - \psi_2^*) \geqslant n - 1$. So $S(\psi_1^* - \psi_2^*) \geqslant n - 1$, and, since n is arbitrary, $S(\psi_1^* - \psi_2^*) = \infty$.

CASE 3. $S = n$, where n is a positive integer. Let x_j ($j = 1, 2, \cdots, n$) be the point at which the jth sign change occurs and let $x_0 = 0$. Without loss of generality, we assume that the sign change at x_n is $+$ to $-$, from which it follows that $b_2 = b \leqslant b_1$. We consider various subcases. (In the remainder of this proof, all intervals (c, d), $(c, d]$, etc., will denote the one point set $\{c\}$ when $c = d$. Also, c may equal ∞.)

CASE 3a. $\psi_1'(W) = \infty$. (Note that W may equal ∞.) Then $W = b_1 = b_2 = b > 0$ and there exist positive real numbers $W_m \uparrow W$ as $m \to \infty$ such that $\psi_1'(W_m) = y_m^* = \psi_2'(W_m) < \infty$. From Theorem 3.1, $S_m = S_{[0, W_m]}(\psi_1 - \psi_2) = S_{[0, y_m^*]}(\psi_1^* - \psi_2^*) = S_m^*$. Also, as $m \to \infty$, $S_m \to S_{[0, W)}(\psi_1 - \psi_2)$ and $S_m^* \to S_{[0, \infty)}(\psi_1^* - \psi_2^*)$ since $y_m^* \uparrow \infty$. So,

$$S_{[0, \infty)}(\psi_1^* - \psi_2^*) = S_{[0, W)}(\psi_1 - \psi_2). \tag{3.5}$$

Now, $\psi_1(x) = \infty = \psi_2(x)$ for $x > W$ since $\psi_1'(x) = \infty = \psi_2'(x)$ for $x \geqslant W$; so,

$$S_{(W, \infty)}(\psi_1 - \psi_2) = 0. \tag{3.6}$$

Furthermore, since $\psi_1(x) = \infty = \psi_2(x)$ for $x > W$, it follows from Definition 2.2 that $\psi_1 - \psi_2$ cannot have a sign change at W. Combining this with (3.5) and (3.6) we get that $S(\psi_1^* - \psi_2^*) = S$.

CASE 3b. $\psi_1'(W) < \infty$. It is clear that $W \in (x_{n-2}, b]$ if $S = n \geqslant 2$, while $W \in [0, b]$ if $n = 1$. The proof concerning the behavior of $S(\psi_1^* - \psi_2^*)$ will depend on the position of W in these intervals. We consider the various cases below.

CASE A. $W = b$. Then either $x_n = W < \infty$ or $x_n < W \leqslant \infty$. If $x_n = W$, then $x_n = b$. So it follows that $W = b = b_2 < b_1$ since a $+$ to $-$ sign change occurs at x_n. Thus $\psi_2'(x) = \infty$ if $x > W$ and there exists an $x' > W$ such that $\psi_1'(x) < \infty$ for $x \leqslant x'$. Thus, since $\psi_1'(W) = y^* = \psi_2'(W)$, it follows from Theorem D that

$$\psi_2^{*\prime}(y) = W < \psi_1^{*\prime}(y) \qquad \text{for} \quad y > y^*. \tag{3.7}$$

Also, from Theorem 3.1, $S_{[0, W]}(\psi_1 - \psi_2) = S - 1 = S_{[0, y^*]}(\psi_1^* - \psi_2^*)$ with $FS_{[0, y^*]}(\psi_1^* - \psi_2^*) = -$. Combining this with (3.7), we see that $\psi_1^* - \psi_2^*$ has exactly one more sign change. So, $S_{[0, \infty)}(\psi_1^* - \psi_2^*) = S$.

If $x_n < W < \infty$, then $S_{[0, W]}(\psi_1 - \psi_2) = S$ and so,

$$S_{[0, y^*]}(\psi_1^* - \psi_2^*) = S \quad \text{and} \quad FS_{[0, y^*]}(\psi_1^* - \psi_2^*) = +, \tag{3.8}$$

by Theorem 3.1. Since $\psi_2(x) = \infty$ for $x > W$, $\psi_2'(x) = \infty$ for $x > W$. Thus, by Theorem D,

$$\psi_2^{*\prime}(y) = W \leqslant \psi_1^{*\prime}(y) \qquad \text{for} \quad y \geqslant y^*, \tag{3.9}$$

since $\psi_1'(W) = y^* = \psi_2'(W)$. Thus, from (3.8) and (3.9), $\psi_1^* - \psi_2^*$ can have no further sign changes on $[y^*, \infty)$, and so, $S(\psi_1^* - \psi_2^*) = S$.

Finally, if $W = \infty$, then there exist positive real numbers $W_m \uparrow \infty$ as $m \to \infty$ such that $\psi_1'(W_m) = y_m^* = \psi_2'(W_m) < \infty$. By Theorem 3.1,

$$S_{[0, y^*)}(\psi_1^* - \psi_2^*) = \lim_m S_{[0, y_m^*]}(\psi_1^* - \psi_2^*) = S. \tag{3.10}$$

Furthermore, since $\psi_i (i = 1, 2)$ is strictly convex on $[0, W)$, it follows from Theorem D that, for $y < y^*$,

$$\psi_i^{*\prime}(y) = \psi_i^{\prime -1}(y) \to W = \infty \qquad \text{as} \quad y \to y^*.$$

So,

$$\psi_i^*(y) = \infty \qquad \text{for} \quad y > y^*. \tag{3.11}$$

Also, $\psi_i^*(y) \to \psi_i^*(y^*)$ as $y \uparrow y^*$ by Theorem D. This with (3.11) shows that $\psi_1^* - \psi_2^*$ cannot have another sign change. So, $S(\psi_1^* - \psi_2^*) = S$ from (3.10).

Case B. $W \in (x_n, b)$. Because of Case A we may assume that $x_n < b$. Since $W \in (x_n, b)$, $S_{[0, y^*]}(\psi_1^* - \psi_2^*) = S$ and $FS_{[0, y^*]}(\psi_1^* - \psi_2^*) = +$. Now either

$$\psi_1'(x) < \psi_2'(x) \qquad \text{for all} \quad x \in (W, b) \tag{3.12}$$

or

$$\psi_1'(x) > \psi_2'(x) \qquad \text{for all} \quad x \in (W, b). \tag{3.13}$$

If (3.12) holds, then by Theorem D and (iii),

$$\psi_1^{*\prime}(y) \geqslant \psi_2^{*\prime}(y) \qquad \text{on} \quad [y^*, \infty).$$

Hence $\psi_1^* - \psi_2^*$ can have no further sign changes.

If (3.13) holds, then, for each $y \in (y^*, \psi_2'(b))$, there exists an $x \in (W, b)$ such that $y = \psi_2'(x)$. Thus,

$$\begin{aligned} \psi_1^*(y) - \psi_2^*(y) &= \psi_1^*(y) - (xy - \psi_2(x)) \\ &= \psi_1^*(y) - (xy - \psi_1(x)) + (\psi_2(x) - \psi_1(x)) \\ &\geqslant 0 \qquad \text{for each} \quad y \in (y^*, \psi_2'(b)), \end{aligned} \tag{3.14}$$

since $x \geqslant x_n$ and $\psi_1^*(y) \geqslant (xy - \psi_1(x))$. Thus, if $\psi_2'(b) = \infty$, it follows from (3.14) that

$$\psi_1^*(y) \geqslant \psi_2^*(y) \qquad \text{for} \quad y \geqslant y^*. \tag{3.15}$$

Now let $\psi_2'(b) < \infty$. We first show that $b \neq \infty$. Suppose $b = \infty$. Then,

$$\psi_1'(b) = \psi_2'(b) < \infty, \tag{3.16}$$

since otherwise it would follow from (iii) and (3.13) that $\psi_1(x) > \psi_2(x)$ for all sufficiently large x which contradicts $FS(\psi_1 - \psi_2) = -$. But (3.16) implies that $W = b$ which contradicts $W \in (x_n, b)$. So $b \neq \infty$.

Finally, if $b < \infty$, then from Theorem D,

$$\psi_2^{*\prime}(y) = b \leqslant \psi_1^{*\prime}(y) \qquad \text{for} \quad y > \psi_1'(b), \tag{3.17}$$

where we ignore (3.17) if $\psi_1'(b) = \infty$. Thus, if we show that $\psi_1^*(y) \geqslant \psi_2^*(y)$ for all $y \in (\psi_2'(b), \psi_1'(b))$ whenever $\psi_2'(b) < \psi_1'(b)$, then this together with (3.14), (3.17) and Theorem C will show that (3.15) holds.

Let $\psi_1'(b) > \psi_2'(b)$ and let $y \in (\psi_2'(b), \psi_1'(b))$. Note that since $\psi_2'(b) < \infty$ and $b < \infty$, $\psi_2(b) = \int_0^b \psi_2'(x)dx < \infty$. So $\psi_1(b) < \infty$ since otherwise $FS(\psi_1 - \psi_2) = +$ which is a contradiction.

Thus, for $x = \psi_1'^{-1}(y)$,

$$\begin{aligned}(3.18)\quad \psi_1^*(y) - \psi_2^*(y) &= yx - \psi_1(x) - (yb - \psi_2(b))\\ &= \psi_2(b) - \psi_1(b) + \psi_1(b) - \psi_1(x) + y(x - b)\\ &= \psi_2(b) - \psi_1(b) + \int_0^b(\psi_1'(z) - y)dz,\end{aligned}$$

by Theorems C and D. Since $FS(\psi_1 - \psi_2) = -$, $\psi_2(b) \geqslant \psi_1(b)$ and since $\psi_1'(x) = y$, $\psi_1'(z) \geqslant y$ for $z \geqslant x$. Thus the last term in (3.18) is nonnegative from which it follows that (3.15) holds.

In view of (3.15), it follows that $\psi_1^* - \psi_2^*$ can have no further sign changes on $[y^*, \infty)$. So, $S(\psi_1^* - \psi_2^*) = S$ for all the cases considered in Case B.

CASE C. $W = x_n$. Again, because of Case A, we may assume that $x_n < b$. It follows from Theorem 3.1 that

$$(3.19)\quad S_{[0,y^*]}(\psi_1^* - \psi_2^*) = S - 1 \quad \text{and} \quad FS_{[0,y^*]}(\psi_1^* - \psi_2^*) = -.$$

Since $W = x_n < b$ and a + to − sign change occurs at x_n it follows that (3.12) must hold. Thus,

$$(3.20)\quad \psi_1^{*\prime}(y) > \psi_2^{*\prime}(y) \quad \text{for} \quad y \in (y^*, \psi_2'(b))$$

and, in fact,

$$(3.21)\quad \psi_1^{*\prime}(y) \geqslant \psi_2^{*\prime}(y) \quad \text{for} \quad y \geqslant y^*.$$

Since $W = x_n$,

$$\psi_1^*(y^*) - \psi_2^*(y^*) = \psi_2(x_n) - \psi_1(x_n) = 0,$$

it follows from (3.20) and (3.21) that

$$\psi_1^*(y) \geqslant \psi_2^*(y) \quad \text{for} \quad y \geqslant y^*,$$

with the inequality strict for $y \in (y^*, \psi_2'(b))$. Combining this with (3.19), we see that $S(\psi_1^* - \psi_2^*) = S$.

CASE D. $W \in (x_{n-1}, x_n)$. From Theorem 3.1, $S_{[0,y^*]}(\psi_1^* - \psi_2^*) = S - 1$ and $FS_{[0,y^*]}(\psi_1^* - \psi_2^*) = -$. If $b > x_n$, then (3.12) must hold. Thus, from Theorem D and (iii),

$$(3.22)\quad \psi_1^{*\prime}(y) \geqslant \psi_2^{*\prime}(y) \quad \text{for} \quad y \geqslant y^*.$$

Also, for $y' = \psi_2'(x_n)$,

$$\begin{aligned}(3.23)\quad \psi_2^*(y') &= x_n y' - \psi_2(x_n)\\ &= x_n y' - \psi_1(x_n)\\ &< \psi_1^*(y'),\end{aligned}$$

where the strict inequality (rather than a weak inequality) follows since ψ_1^* is strictly convex and $\psi_1'(x_n) < \psi_2'(x_n)$. Thus, $S(\psi_1^* - \psi_2^*) = S$ by (3.22) and (3.23).

If $b = x_n < \infty$, then $b < b_1$ and $\psi_1(x) < \infty = \psi_2(x)$ for $x \in (x_n, b_1)$. Thus, from properties of convex functions and their conjugates (see Section 15 of Roberts and

Varberg (1973)),

(3.24) $$\psi_2^{*\prime}(y) \leqslant b < \psi_1^{*\prime}(y) \quad \text{for} \quad y > \psi_1'(b).$$

Also, since either (3.12) or (3.13) holds, either

(3.25) $$\psi_1^{*\prime}(y) > \psi_2^{*\prime}(y) \quad \text{for} \quad y \in (y^*, \psi_1'(b))$$

or

(3.26) $$\psi_1^{*\prime}(y) < \psi_2^{*\prime}(y) \quad \text{for} \quad y \in (y^*, \psi_1'(b)),$$

by Theorem D. Furthermore, since $b_2 = b = x_n < \infty$, $\psi_2^*(y) < \infty$ for $y \geqslant 0$. Thus,

(3.27) $$\psi_1^*(y) - \psi_2^*(y) = \textstyle\int_0^y (\psi_1^{*\prime}(z) - \psi_2^{*\prime}(z))dz,$$

whenever $\psi_1^*(y) < \infty$ by Theorem C. Combining this with (3.24), (3.25) and (3.26), we see that $\psi_1^* - \psi_2^*$ can have had at most one more sign change on $[y^*, \infty)$, and since $FS_{[0,y^*]}(\psi_1^* - \psi_2^*) = -$, it must have one more sign change on $[y^*, \infty)$ because of (3.24), (3.27) and the monotonicity of $\psi_1^{*\prime}$. Thus $S(\psi_1^* - \psi_2^*) = S$.

CASE E. $n = 1$ and $W = 0$. Since $W = 0$ and $\psi_1 - \psi_2$ has a $+$ to $-$ sign change, it follows from the definition of W that

(3.28) $$\psi_1'(x) > \psi_2'(x) \quad \text{for} \quad x \in (0, b).$$

Since $\psi_1 - \psi_2$ has a $+$ to $-$ sign change it is easy to see that the following must hold because of (3.28). First, $b < \infty$ and $x_1 = b$; second, $\psi_1'(b) > \psi_2'(b)$ (from definition of W and (iii)); third, $\psi_2'(x) = \infty$ for $x > b$; and fourth, there is an $\varepsilon > 0$ such that $\psi_1'(x) < \infty$ for $x \in [b, b + \varepsilon)$. From these four facts, (3.28) and Theorem D, it readily follows that

(3.29) $$\begin{array}{ll} \psi_1^{*\prime}(y) < \psi_2^{*\prime}(y) & \text{for} \quad y \in (0, \psi_2'(b)] \\ \psi_1^{*\prime}(y) \leqslant b = \psi_2^{*\prime}(y) & \text{for} \quad y \in (\psi_2'(b), \psi_1'(b)] \\ \psi_1^{*\prime}(y) > b = \psi_2^{*\prime}(y) & \text{for} \quad y > \psi_1'(b). \end{array}$$

Thus $S(\psi_1^* - \psi_2^*) = 1$ from (3.29) and Theorem C.

CASE F. $n \geqslant 2$ and $W \in (x_{n-2}, x_{n-1}]$. From Theorem 3.1, $S_{[0,y^*]}(\psi_1^* - \psi_2^*) = S - 2$ and $FS_{[0,y^*]}(\psi_1^* - \psi_2^*) = +$. Since $W \in (x_{n-2}, x_{n-1}]$ and $\psi_1 - \psi_2$ has a $-$ to $+$ sign change at x_{n-1}, (3.13) must hold. Thus, since $\psi_1 - \psi_2$ has a $+$ to $-$ sign change at x_n, it follows from (3.13) that $x_n = b_2 = b < b_1, y' = \psi_1'(b) < \infty$, and

(3.30) $$\psi_1^{*\prime}(y) < \psi_2^{*\prime}(y) \quad \text{for} \quad y \in (y^*, y')$$

and

(3.31) $$\psi_2^{*\prime}(y) = b < \psi_1^{*\prime}(y) \quad \text{for} \quad y > y',$$

by Theorem D. Thus, via an argument similar to the one following (3.27), it follows from (3.30) and (3.31) that $\psi_1^* - \psi_2^*$ has either no more sign changes or exactly two more sign changes on $[y^*, \infty)$. Thus $S(\psi_1^* - \psi_2^*)$ is either equal to $S - 2$ or S.

This exhausts all the possible cases and proves the part of Theorem 3.2 concerning the number of sign changes of $\psi_1^* - \psi_2^*$. Because of Theorem 3.1, it is clear from the above proof that the sequence of signs of $\psi_1^* - \psi_2^*$ behaves as stated in Theorem 3.2. This completes the proof of Theorem 3.2.

4. Behavior of the differences of large deviation rates. Let $X_1, X_2, \cdots$ be i.i.d. random variables with distribution $F_1(F_2)$ such that $\int x dF_1(x)(\int x dF_2(x)) = 0$. To avoid trivialities, it is assumed that F_1 and F_2 are not degenerate and $F_1 \neq F_2$. If the densities of F_1 and F_2 both exist with respect to some σ-finite measure μ, they will be denoted by f_1 and f_2, respectively.

For $i = 1$ and 2, we define the following quantities. Let

$$\phi_i(\lambda) = \int \exp(\lambda x) dF_i(x) \qquad \text{and} \qquad \psi_i(\lambda) = \log \phi_i(\lambda),$$

i.e., ϕ_i is the m.g.f. of F_i and ψ_i is the c.g.f. of F_i. Then it follows from (1.1) that, for $i = 1$ and 2,

$$\lim n^{-1} \log P_{F_i}(\bar{X}_n \geqslant a) = -\psi_i^*(a) \qquad \text{for} \quad a \geqslant 0,$$

where

$$\psi_i^*(a) = \sup\{\lambda a - \psi_i(\lambda): \lambda \geqslant 0\}$$

and

$$\lim n^{-1} \log P_{F_i}(\bar{X}_n \leqslant a) = -\psi_i^*(a) \qquad \text{for} \quad a \leqslant 0,$$

where

$$\psi_i^*(a) = \sup\{\lambda a - \psi_i(\lambda): \lambda \leqslant 0\}.$$

In this section, we shall study how the behavior of the sign changes of $\psi_1^* - \psi_2^*$ is related to that of $\psi_1 - \psi_2$ (or equivalently, $\phi_1 - \phi_2$), $F_1 - F_2$, and $f_1 - f_2$. To do this, we shall first need the following theorem which enumerates some well known properties of ψ_1 and ψ_2.

For $i = 1$ and 2, let $b_i = \sup\{\lambda: \phi_i(\lambda) < \infty\}$ and let $b = \min\{b_1, b_2\}$.

THEOREM 4.1. *If $b > 0$, then for $i = 1$ and 2,*

(a) *ψ_i is strictly convex and strictly increasing on $[0, b_i)$ with $\psi_i(0) = 0$,*
(b) *$\psi_i(\lambda) \to \psi_i(b_i)$ as $\lambda \uparrow b_i$,*
(c) *ψ_i' exists and is continuous on $[0, b_i)$ with $\psi_i'(0) = 0$ and $\psi_i'(\lambda) \uparrow \psi_i'(b_i)$ as $\lambda \uparrow b_i$, and*
(d) *$\psi_1 - \psi_2$ is not identically zero on any subinterval of $[0, b)$.*

From Theorem 4.1, we have the following theorem as an immediate consequence of Theorem 3.2.

THEOREM 4.2. *For ψ_1, ψ_2, ψ_1^*, and ψ_2^* as defined above, the relationship of the sign changes of $\psi_1^* - \psi_2^*$ to that of $\psi_1 - \psi_2$ is as indicated in Theorem 3.2.*

LEMMA 4.3. *For $I = (-\infty, \infty)$,*

$$S_I(\psi_1^* - \psi_2^*) \leqslant S_I(F_1 - F_2) \leqslant S_I(f_1 - f_2). \tag{4.1}$$

PROOF. If f_1 and f_2 exist, then, since $F_1(x) - F_2(x) = \int_{-\infty}^{x}[f_1(y) - f_2(y)]d\mu(y)$, $S(F_1 - F_2) \leqslant S(f_1 - f_2)$. This proves the right side of (4.1).

To prove the left side, note that, for the purpose of sign changes, we need only consider when either $\phi_1(\lambda)$ or $\phi_2(\lambda)$ is finite. When this is the case,

$$\phi_1(\lambda) - \phi_2(\lambda) = \lambda\int_{-\infty}^{\infty} \exp(\lambda x)[F_2(x) - F_1(x)]dx. \tag{4.2}$$

For each $K > 0$, let

$$g_K(\lambda) = \int_{-K}^{K} \exp(\lambda x)[F_2(x) - F_1(x)]dx.$$

Since $\exp(\lambda x)$ is TP (see Example 2.1), it follows from the VD Theorem that

$$S(g_K) \leqslant S(F_1 - F_2).$$

Since $g_K(\lambda) \to \phi_1(\lambda) - \phi_2(\lambda)$ as $K \to \infty$ whenever $\phi_1(\lambda)$ or $\phi_2(\lambda)$ is finite, it follows that $S(\psi_1 - \psi_2) = S(\phi_1 - \phi_2) \leqslant S(F_1 - F_2)$. Thus the left side of (4.1) will follow from Theorem 4.2 by observing that $\psi_1^* - \psi_2^*$ has a sign change at 0 if and only if $\psi_1 - \psi_2$ does.

As is exemplified in Example 4.6 below, the bounds in (4.1) may not be very good, and as such may not be very useful in determining the behavior of $S(\psi_1^* - \psi_2^*)$. However, if $F_1 - F_2$ or $f_1 - f_2$ have very few sign changes, then $S(\psi_1^* - \psi_2^*)$ may be totally determined by the behavior of $F_1 - F_2$ or of $f_1 - f_2$. These results are given in the following lemma and its corollary.

LEMMA 4.4. *Let*

$$1 \leqslant S_{(-\infty,\infty)}(F_1 - F_2) \leqslant 2 \quad \text{and} \quad FS_{(-\infty,\infty)}(F_1 - F_2) = -. \tag{4.3}$$

Then,

$$\psi_2^*(a) \geqslant \psi_1^*(a) \quad \text{for} \quad a \geqslant 0. \tag{4.4}$$

PROOF. If we show that

$$\psi_1(\lambda) \geqslant \psi_2(\lambda) \quad \text{for} \quad \lambda \geqslant 0, \tag{4.5}$$

then (4.4) will immediately follow.

If for $\lambda > 0$, $\psi_2(\lambda) = \infty$, then $\psi_1(\lambda) = \infty$ since $FS(F_1 - F_2) = -$. Hence, if $\psi_2(\lambda) = \infty$ for all $\lambda > 0$, (4.5) trivially holds. Thus, assume that there exists a $\delta > 0$ such that

$$\psi_1(\lambda) < \infty \quad \text{for} \quad 0 \leqslant \lambda \leqslant \delta. \tag{4.6}$$

For $\sigma > 0$ and $i = 1$ and 2, let $F_{i\sigma}$ denote the convolution of F_i with a normal distribution having mean 0 and variance σ^2. It follows from the VD Theorem (via an argument similar to that following (4.2)) that

$$S(F_{1\sigma} - F_{2\sigma}) \leqslant S(F_1 - F_2);$$

and since $\lim_{\sigma\to 0} (F_{1\sigma}(x) - F_{2\sigma}(x)) = F_1(x) - F_2(x)$ if x is a continuity point of $F_1 - F_2$ and since $S(F_1 - F_2) < \infty$, it follows that

$$S(F_{1\sigma} - F_{2\sigma}) = S(F_1 - F_2) \tag{4.7}$$

for all sufficiently small σ, say $0 < \sigma < \sigma_0$.

Fix $\sigma \in (0, \sigma_0)$ and let Y_1 and Y_2 denote two random variables with distributions $F_{1\sigma}$ and $F_{2\sigma}$, respectively. For $i = 1$ and 2, let

$$\begin{aligned} Y_i(K) &= Y_i \quad \text{if} \quad Y_i > K \\ &= K \quad \text{if} \quad Y_i \leqslant K. \end{aligned}$$

It follows from (4.6) that, for each $K > -\infty$, the c.g.f.'s of $Y_1(K)$ and $Y_2(K)$, ψ_{1K} and ψ_{2K}, respectively, are finite in an open interval about 0.

Since $E(Y_1(-\infty)) = 0 = E(Y_2(-\infty))$ and $F_{i\sigma}$ has support $(-\infty, \infty)$ for $i = 1$ and 2, it follows that there exist two sequences $\{K(n)\}$ $\{M(n)\}$ with $K(n) < M(n) \to -\infty$ as $n \to \infty$ such that

$$E(Y_1(M(n)) - Y_2(K(n))) > 0. \tag{4.8}$$

Since the left side of (4.8) is the derivative of $\psi_{1, M(n)} - \psi_{2, K(n)}$ at 0 and $\psi_{1, M(n)}(0) - \psi_{2, K(n)}(0) = 0$, it follows from (4.8) that $\psi_{1, M(n)} - \psi_{2, K(n)}$ has a $-$ to $+$ sign change at 0. Let $G_1(G_2)$ denote the distribution of $Y_1(M(n))(Y_2(K(n)))$ for a fixed n. It will follow from (4.3), (4.7), and the construction of G_1 and G_2 that

$$S(G_1 - G_2) = 2$$

for n sufficiently large and the sequence of signs of $G_1 - G_2$ is $-, +, -$. Replacing $F_2 - F_1$ with $G_2 - G_1$ in (4.2), it thus follows from the VD Theorem that $S(\psi_{1, M(n)} - \psi_{2, K(n)}) \leqslant 2$; and if $S(\psi_{1, M(n)} - \psi_{2, K(n)}) = 2$, the sequence of signs of $\psi_{1, M(n)} - \psi_{2, K(n)}$ is $+, -, +$. Thus, since $\psi_{1, M(n)} - \psi_{2, K(n)}$ has a $-$ to $+$ sign change at 0,

$$\psi_{1, M(n)}(\lambda) \geqslant \psi_{2, K(n)}(\lambda) \qquad \text{for} \quad \lambda \geqslant 0.$$

Letting $n \to -\infty$, it follows that

$$\int \exp(\lambda x) dF_{1\sigma}(x) \geqslant \int \exp(\lambda x) dF_{2\sigma}(x) \qquad \text{for} \quad \lambda \geqslant 0,$$

from which (4.5) follows by letting $\sigma \to 0$.

COROLLARY 4.5. *If*

$$S_{(-\infty, \infty)}(f_1 - f_2) = 2$$

and the last sign change is from $-$ *to* $+$, *then* (4.4) *holds.*

PROOF. It is easy to see that $S_{(-\infty, \infty)}(F_1 - F_2) = 1$ and the sign change is from $+$ to $-$. Result then follows from Lemma 4.4.

EXAMPLE 4.6. Let F_1 and F_2 be two symmetric distributions, i.e., $F_i(-x) = 1 - F_i(x)$ for $i = 1$ and 2. For $k = 1, 2, \cdots$, and $i = 1$ and 2, let

$$\mu_{k, i} = \int_{-\infty}^{\infty} x^{2k} dF_i(x) = 2k \int_0^{\infty} x^{2k-1} \bar{F}_i(x)\, dx, \tag{4.9}$$

where $\bar{F}_i(x) = 1 - F(x)$. For $k = 1, 2, \cdots, n$, let

$$\mu_{k, 1} = \mu_{k, 2} < \infty \tag{4.10}$$

and for $k > n$,

$$\mu_{k, 1} \leqslant \mu_{k, 2}.$$

Then it follows that

$$\phi_1(\lambda) \leqslant \phi_2(\lambda) \qquad \text{for every} \quad \lambda.$$

Hence, from Theorem 4.2,

$$S(\psi_1^* - \psi_2^*) = 0 \qquad \text{and} \quad \psi_1^*(a) \geqslant \psi_2^*(a)$$

for all a. However, since x^k is a strictly totally positive function for $x \geqslant 0$ and $k = 1, 2, \cdots$, if follows from (4.9), (4.10) and the VD Theorem that $S(F_1 - F_2) \geqslant n$. (Here, we are appealing to a more general version of the VD Theorem than the one which appears in Section 2. See Karlin (1968), page 21, (3.5).)

REMARK 4.7. In this paper, the only statistic considered was $\overline{X}_n$. However, in many instances (see, e.g., Sievers (1969) and Bahadur (1971)), a statistic T_n will satisfy,

$$n^{-1} \log P_i(T_n \geqslant a) \to -\psi_i^*(a) \quad \text{for} \quad i = 1 \quad \text{and} \quad 2,$$

where $\psi_i^*(a) = \sup\{\lambda a - \psi_i(\lambda) : \lambda \geqslant 0\}$ and $n^{-1} \log E_i(\exp(\lambda T_n)) \to \psi_i(\lambda)$. If ψ_1 and ψ_2 satisfy (i), (ii), (iii) and (iv′), it is clear that the conclusions of Theorem 4.2 would hold for this case.

Acknowledgment. The author wishes to thank J. Sethuraman for the seed of the problem from which this paper grew and would also like to thank the referee whose many helpful comments improved this paper.

REFERENCES

[1] BAHADUR, R. R. (1971). *Some Limit Theorems in Statistics.* Regional Conference Series in Applied Mathematics. *SIAM*, Philadelphia.

[2] CHERNOFF, H. (1952). A measure of asymptotic efficiency of tests of hypothesis based on the sum of observations. *Ann. Math. Statist.* **23** 493–507.

[3] KARLIN, S. (1968). *Total Positivity.* Stanford Univ. Press.

[4] ROBERTS, A. and VARBERG, D. (1973). *Convex Functions.* Academic Press, New York.

[5] SIEVERS, G. L. (1969). On the probability of large deviations and exact slopes. *Ann. Math. Statist.* **40** 1908–1921.

DEPARTMENT OF STATISTICS
219 POND LABORATORY
PENNSYLVANIA STATE UNIVERSITY
UNIVERSITY PARK, PENNSYLVANIA 16802