stichting
mathematisch
centrum

MC

AFDELING MATHEMATISCHE STATISTIEK SW 79/81 DECEMBER
(DEPARTMENT OF MATHEMATICAL STATISTICS)

C.A.J. KLAASSEN

A NOTE ON AN INEQUALITY OF CHERNOFF

Preprint

kruislaan 413 1098 SJ amsterdam

Printed at the Mathematical Centre, 413 Kruislaan, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

1980 Mathematics subject classification: Primary : 60E05
Secondary : 26D10

A note on an inequality of Chernoff*)

by

Chris A.J. Klaassen

ABSTRACT

An inequality due to Chernoff is generalized.

KEY WORDS & PHRASES: *Inequality of Chernoff, Cauchy-Schwarz inequality*

---

*) This report will be submitted for publication elsewhere.

1. INTRODUCTION

Let X be a standard normal random variable and let $G : \mathbb{R} \to \mathbb{R}$ be a function which is absolutely continuous with respect to Lebesgue measure with Radon-Nikodym derivative g. In CHERNOFF (1980,1981) the elegant inequality

$$\text{(1.1)} \qquad \operatorname{var} G(X) \le E\, g^2(X)$$

has been presented. For an arbitrary random variable X we shall derive the generalization (2.7) of (1.1). Although this derivation has not been restricted to the normal case, it seems to be somewhat simpler than the one given in CHERNOFF (1981). The main idea in the proof of (2.7), namely the use of the Cauchy-Schwarz inequality, is also contained in CACOULLOS (1981), which restricts attention to the case $h = 1$, $\alpha = c = 0$ in Theorem 2.1, and in CHEN (1980), which considers the multivariate normal case.

2. THE RESULT AND SOME EXAMPLES

For our formulation of Chernoff's inequality we need the convention that the variance of a random variable is infinite iff the second moment of that random variable is infinite. Furthermore, for $a \le b$ we'll denote an integral over $(a,b]$ by $\int_a^b$ or $-\int_b^a$.

<u>THEOREM 2.1</u>. *Let* μ *be a* σ*-finite measure on* $(\mathbb{R}, \mathcal{B})$ *and let* X *be a random variable with density* f *with respect to* μ. S *is the smallest interval containing the support of* f *and* $h : S \to \mathbb{R}$ *is a nonnegative measurable function such that for all* $a,b \in S$

$$\text{(2.1)} \qquad \left| \int_a^b h\,d\mu \right| < \infty$$

and

$$\text{(2.2)} \qquad E_f \left| \int_a^X h\,d\mu \right| < \infty.$$

Now there exist c ∈ S *and* α ∈ [0,1] *such that*

$$E_f\, H(X) = 0 \tag{2.3}$$

holds for

$$H(x) = \int_c^x h\,d\mu + \alpha h(c)\mu(\{c\}), \qquad x \in S. \tag{2.4}$$

Let $g : S \to \mathbb{R}$ *be a measurable function satisfying* (2.1) *and for some* $d \in S$ *and* $e \in \mathbb{R}$ *define*

$$G(x) = \int_d^x g\,d\mu + e, \qquad x \in S. \tag{2.5}$$

If

$$\mu(\{x \in S \mid g(x) \neq 0,\ f(x)h(x) = 0\}) = 0, \tag{2.6}$$

then the inequality

$$\operatorname{var}_f G(X) \le E_f\left\{\frac{g^2(x)}{f(x)h(x)} \int_{S\cap[X,\infty)} Hf\,d\mu\right\} \tag{2.7}$$

is valid, with equality iff the variance is infinite or g *is a multiple of* h μ *- almost everywhere on* S.

The proof of this theorem will be presented in section 3. In Table 2.1 the upperbound B from (2.7) is given for some choices of f and H with μ Lebesgue measure. Note that (1.1) can be obtained from (2.7) by choosing $h(x) = 1$, $x \in \mathbb{R}$ (cf. examples 1 and 6 of Table 2.1).

name	f	H	B
1 normal	$(2\pi\sigma^2)^{-\frac{1}{2}}\ e^{-\frac{1}{2}x^2\sigma^{-2}}$	x	$\sigma^2\ E_f\ g^2(X)$
2 exponential	$\lambda e^{-\lambda x}, \quad x > 0$	$(x-\frac{2}{\lambda})e^{\frac{1}{2}\lambda x}$	$4\lambda^{-2}E_f\ g^2(X)$
3 Laplace	$\frac{1}{2}\lambda e^{-\lambda\lvert x\rvert}$	$x\ e^{\frac{1}{2}\lambda\lvert x\rvert}$	$4\lambda^{-2}E_f\ g^2(X)$
4 logistic	$2\lambda(e^{\lambda x}+e^{-\lambda x})^{-2}$	$e^{\lambda x}-e^{-\lambda x}$	$\lambda^{-2}\ E_f\ g^2(X)$
5 gamma	$[\sigma^{\alpha}\Gamma(\alpha)]^{-1}\ x^{\alpha-1}\ e^{-\frac{x}{\sigma}}, x>0$	$x-\alpha\sigma$	$\sigma\ E_f\ X\ g^2(X)$
6 -	$c(\alpha,\sigma)\lvert x\rvert^{\alpha-1}\ e^{-\frac{1}{2}\lvert\frac{x}{\sigma}\rvert^{\alpha+1}}$	x	$\frac{2\sigma^{\alpha+1}}{\alpha+1}\ E_f\lvert X\rvert^{1-\alpha}g^2(X)$

Table 2.1. The value of the upperbound B of (2.7) for some choices of f and H with μ Lebesgue measure.

For μ counting measure on the integers a few examples of (2.7) are given in Table 2.2. We only note the folllowing. If X has a discrete distribution with mean ν then $H(x) = x-\nu$, $x \in \mathbb{Z}$, is realized by the choices $h(x) = 1$, $c = [\nu]+1$ and $\alpha = [\nu]+1-\nu$, where $[\nu]$ denotes the integer part of ν (cf. (2.3) and (2.4)).

name	f	H	B
1 Poisson	$e^{-\lambda}\lambda^{x}(x!)^{-1}$	$x-\lambda$	$E_f X g^2(X)$
2 binomial	$\binom{n}{x}\ p^{x}(1-p)^{n-x}$	$x-np$	$(1-p)E_f X g^2(X)$
3 negative binomial	$\binom{x-1}{k-1}p^{k}(1-p)^{x-k}$	$x-\frac{k}{p}$	$p^{-1}\ E_f(X-k)g^2(X)$

Table 2.2. The value of the upperbound B of (2.7) for some choices of f and H with μ counting measure on the integers.

3. PROOF

If h vanishes μ-almost everywhere on S then (2.3) and (2.4) are fulfilled for all $c \in S$ and $\alpha \in [0,1]$. Therefore, for the proof of (2.3) we assume without loss of generality that h does not have this property. In view of (2.1) and (2.2) we can define $\psi : S \to \mathbb{R}$ by

$$\psi(a) = E_f \int_a^X h d\mu.$$

Since h is nonnegative ψ is nonincreasing on S. Because $\psi(a)$ is finite, Fubini's theorem yields

$$\psi(a) = - \int_{S\cap(-\infty,a]} \int_{S\cap(-\infty,y)} f(x)d\mu(x)h(y)d\mu(y) + \int_{S\cap(a,\infty)} \int_{S\cap[y,\infty)} f(x)d\mu(x)h(y)d\mu(y). \tag{3.1}$$

Let $a_0 = \inf S$. If $a_0 \in S$ we obtain the nonnegativity of $\psi(a_0)$ from (3.1). If $a_0 \notin S$, then we see by the dominated and the monotone convergence theorem that

$$\lim_{a \downarrow a_0} \psi(a) = \int_S \int_{S\cap[y,\infty)} f(x)d\mu(x)h(y)d\mu(y) \geq 0.$$

In both cases there exists a $c_0 \in S$ with $\psi(c_0) \geq 0$. Analogously we see that there exists a $c_1 \in S$ with $\psi(c_1) \leq 0$. If there is a $c \in [c_0,c_1]$ with $\psi(c) = 0$, then (2.3) and (2.4) are valid with $\alpha = 0$. If not, there exists a $c \in (c_0,c_1]$ such that $0 \leq \psi(c-) < \infty$ and $-\infty < \psi(c) < 0$ and hence also an $\alpha \in (0,1]$ such that

$$\alpha\psi(c-) + (1-\alpha)\psi(c) = 0. \tag{3.2}$$

In view of (2.4) equality (3.2) is the same one as (2.3).

By the Cauchy-Schwarz inequality, Fubini's theorem and (2.3) we obtain for g satisfying (2.5) and (2.6)

$$
\begin{aligned}
\operatorname{var}_f G(X) &\le E_f\Big[\int_c^X \frac{g}{h^{\frac12}}\, h^{\frac12}\, d\mu + \alpha\,\frac{g(c)}{h^{\frac12}(c)}\, h^{\frac12}(c)\,\mu(\{c\})\Big]^2 \\
&\le E_f\Big[\int_c^X \frac{g^2}{h}\, d\mu + \alpha\,\frac{g^2(c)}{h(c)}\,\mu(\{c\})\Big] H(X) \\
&= \int_S \int_c^x \frac{g^2(y)}{h(y)}\, d\mu(y) H(x) f(x)\, d\mu(x) \\
&= -\int_{S\cap(-\infty,c]} \int_{S\cap(-\infty,y)} H(x) f(x)\, d\mu(x)\, \frac{g^2(y)}{h(y)}\, d\mu(y) \\
&\quad + \int_{S\cap(c,\infty)} \int_{S\cap[y,\infty)} H(x) f(x)\, d\mu(x)\, \frac{g^2(y)}{h(y)}\, d\mu(y) \\
&= \int_S \frac{g^2(y)}{f(y)h(y)} \int_{S\cap[y,\infty)} H(x) f(x)\, d\mu(x)\; f(y)\, d\mu(y),
\end{aligned}
$$

which implies (2.7) and thereby the theorem. □

ACKNOWLEDGEMENTS.

My thanks go to Prof. H. Chernoff for pointing at the references CHEN (1980) and CACOULLOS (1981) and to Prof. T. Cacoullos for sending me a copy of his stimulating paper.

REFERENCES

CACOULLOS, T. (1981), *On upper and lower bounds for the variance of a function of a random variable*, to appear.

CHEN, L.H.Y. (1980), *An inequality for the multivariate normal distribution*, M.I.T. Technical Report No. 18.

CHERNOFF, H. (1980), *The identification of an element of a large population in the presence of noise*, Ann. Statist. 8, 1179-1197.

CHERNOFF, H. (1981), *A note on an inequality involving the normal distribution*, Ann. Probab. 9, 533-535.

ONTVANGEN 8 FEB. 1982