Probabilistic inequalities via the convexity method

Villő Csiszár
Eötvös Loránd University, Budapest, Hungary

Joint work with T. F. Móri and G. J. Székely

25th European Meeting of Statisticians, Oslo, 27th July, 2005

The convexity method in probability theory

Setting:

- $(\Omega, \mathcal{F})$: a measurable space
- $\mathcal{P}$: a convex family of probability measures defined on it
- $\mathcal{Q} \subset \mathcal{P}$: a subfamily such that $\mathcal{P} = \text{conv}(\mathcal{Q})$
- $f = (f_1, \ldots, f_n) : \Omega \to \mathbb{R}^n$: a Borel measurable function
- $\mathcal{M}(\mathcal{P}) := \left\{ \int_\Omega f \, dP : P \in \mathcal{P} \right\}$: the set of f-moments of $\mathcal{P}$

Claim 1.: $\mathcal{M}(\mathcal{P})$ coincides with the closed convex hull of the set $\mathcal{M}(\mathcal{Q})$.

Claim 2.: An inequality of the form

$$h_1\Big(Ef_1(X), \ldots, Ef_{n-1}(X)\Big) \le Ef_n(X) \le h_2\Big(Ef_1(X), \ldots, Ef_{n-1}(X)\Big),$$

where h_1 and h_2 are convex, resp. concave $(n-1)$-variate functions, holds for every X with distribution in $\mathcal{P}$, if and only if it holds in $\mathcal{Q}$.

<u>Some previous applications of the method</u>:

- T. F. Móri and G. J. Székely: A note on the background of several Bonferroni–Galambos type inequalities, J. Appl. Probab. 22(1985), pp. 836–843

- T. F. Móri and G. J. Székely: On the correlation between the sample mean and sample variance. In: A. Iványi (Ed.), Second Conference of Program Designers. ELTE, Budapest, 1986, pp. 207–209

Some examples

- $\mathcal{P}$:= the set of probability distributions concentrated onto $H \subset \mathbb{R}^k$.
  $\mathcal{Q}$:= the set of degenerate distributions δ_x with $x \in H$.

- $\mathcal{P}$:= the set of probability distributions with fixed expectation $\mu \in \mathbb{R}$.
  $\mathcal{Q} := \{P_{xy} : x < \mu < y\} \cup \{\delta_\mu\}$, where $P_{xy} = \dfrac{y-\mu}{y-x}\,\delta_x + \dfrac{\mu-x}{y-x}\,\delta_y$.

- $\mathcal{P}$:= the family of symmetric unimodal distributions.
  $\mathcal{Q}$:= the family of symmetric uniform distributions $U[-x,x]$, $x \geq 0$.

- $\mathcal{P}$:= the family of unimodal distributions with mode μ.
  $\mathcal{Q} := \{U[x,\mu] : x < \mu\} \cup \{U[\mu,x] : \mu \leq x\}$.

Generalization of some moment-type inequalities of F. Qi and T. K. Pogány

Under what conditions on f does the inequality

$$\int_a^b [f(x)]^\alpha dx \geq \left(\int_a^b f(x)dx \right)^\beta$$

hold for fixed α and β? (This question was studied by e.g. F. Qi, S. Mazouzi and F. Qi, K. W. Yu and F. Qi, N. Towghi, T. K. Pogány.)

Consider the interval $[a, b]$, with the uniform distribution, and the random variable $X = f$. Then the previous inequality becomes

$$E(X^\alpha) \geq C\,(EX)^\beta, \text{ where } C = (b-a)^{\beta-1}.$$

Theorem 1. *$H = [x_1, x_2]$ is called a maximal interval of type I if $E(X^\alpha) \geq C\,(EX)^\beta$ holds for every random variable X such that $P(X \in H) = 1$, and no larger interval exists with the same property. Let x_0 be the only positive solution of the equation $x^\alpha = Cx^\beta$, i.e., $x_0 = C^{\frac{1}{\alpha-\beta}}$.*

Suppose $\alpha < 1$. Then all maximal intervals $H = [x_1, x_2]$ of type I can be obtained in the following way: let $z \geq x_0$ be arbitrary, and let $x_1 < x_2$ be the positive solutions of the equation

$$x^\alpha = C(1-\beta)z^\beta + C\beta z^{\beta-1}x.$$

On the other hand, if $1 \leq \alpha$, the only maximal interval of type I is $H = [x_0, +\infty)$.

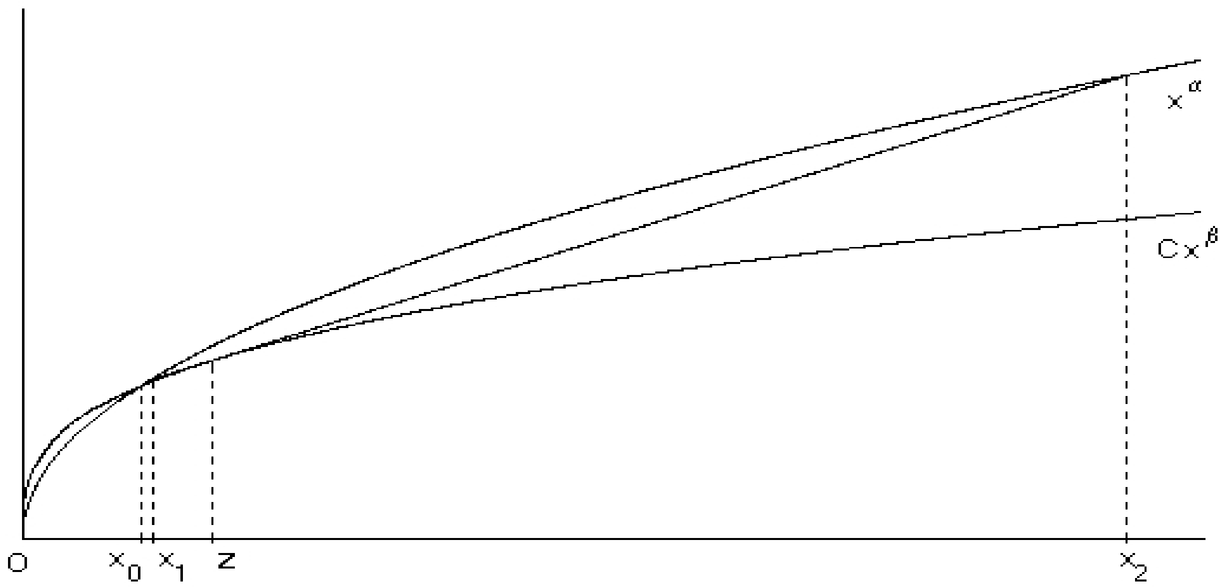
A similar theorem holds for the converse inequality $E(X^\alpha) \leq C\,(EX)^\beta$.

Sketch of proof

Let $H = [x_1, x_2]$ be an interval. The moment set $\mathcal{M} = \Big\{(EX, E(X^\alpha)) : P(X \in H) = 1\Big\}$ is equal to the convex hull of $\Big\{(x, x^\alpha) : x \in H\Big\}$. Now, $E(X^\alpha) \geq C\,(EX)^\beta$ holds if and only if $\mathcal{M}$ lies entirely above the graph of the function $x \mapsto C\,x^\beta$, $x \in H$.

When $\alpha < 1$, the function $x \mapsto x^\alpha$ is concave, hence the moment set over H is just the area between the curve of the function and its chord, and the chord lies below the curve. By the maximality of the interval, the chord must be tangential to the curve $x \mapsto Cx^\beta$. The straight line described in the theorem is just the tangent to the curve $x \mapsto Cx^\beta$ at point $x = z$.

If $\alpha \geq 1$, the function $x \mapsto x^{\alpha}$ is convex, hence the moment set lies above the curve. Since $x^{\alpha} \geq Cx^{\beta}$ for $x \geq x_0$, $H = [x_0, +\infty)$ is a type I interval and it cannot be extended, because $x^{\alpha} < Cx^{\beta}$ for $0 < x < x_0$.



Diaz–Metcalf type inequalities

Diaz–Metcalf inequality (in a probabilistic setting): Let ξ and η be bounded, $P(m_1 \le \xi \le M_1) = 1$, $P(m_2 \le \eta \le M_2) = 1$, $m_2 > 0$. Then

$$m_2 M_2 E\xi^2 + m_1 M_1 E\eta^2 \le (m_1 m_2 + M_1 M_2) E\xi\eta,$$

and equality holds if and only if $P(m_2\xi = M_1\eta$ or $M_2\xi = m_1\eta) = 1$.

Theorem 2. *Let ξ and η be as above, with positive m_1. Let a and b be positive numbers. The smallest positive c such that*

$$aE\xi^2 + bE\eta^2 \le cE\xi\eta$$

always holds is given by

$$c = \begin{cases} a\,\dfrac{m_1}{M_2} + b\,\dfrac{M_2}{m_1}, & \text{if} \quad \dfrac{a}{b} \le \dfrac{m_2 M_2}{m_1 M_1}, \\ a\,\dfrac{M_1}{m_2} + b\,\dfrac{m_2}{M_1}, & \text{if} \quad \dfrac{a}{b} > \dfrac{m_2 M_2}{m_1 M_1}. \end{cases}$$

As a generalization to the problem, let $X = \xi^2$ and $Y = \eta^2$; then $\xi\eta = \sqrt{XY}$, a concave function of X and Y, and $P\Big[(X,Y) \in \mathcal{D}\Big] = 1$, where $\mathcal{D}$ is a rectangle.

Theorem 3. *For any $\varphi : \mathcal{D} \to \mathbb{R}$ concave function, we gave exact lower bounds for $E\varphi(X,Y)$ of the form $pEX + qEY + r$.*

Chebyshev-type inequalities for scale mixtures

Let F be the p.d.f. of a fixed symmetric distribution, and X a mixture of the scale-transforms of the distribution F.

With given $\alpha > 0$, we look for the minimal value of C_α which satisfies for every $t > 0$

$$P(|X| \geq t) \leq C_\alpha \frac{E|X|^\alpha}{t^\alpha}.$$

Theorem 4. *Suppose that F has a density f, which is continuous and positive over a finite or infinite interval, and 0 outside of it. Suppose further, that for $z > 0$, $z^{1+\alpha} f(z)$ is initially increasing, then decreasing. Let z_α be the smallest positive root of the equation*

$$\frac{z f(z)}{\overline{F}(z)} = \alpha.$$

In terms of this z_α and

$$\varrho = \left(\frac{E|X|^\alpha}{M_\alpha} \right)^{1/\alpha}$$

we have

$$\max P(|X| \geq t\varrho) = \begin{cases} 2\,\overline{F}(z_\alpha) z_\alpha^\alpha\, t^{-\alpha}, & \text{if } t \geq z_\alpha, \\ \\ 2\,\overline{F}(t) & \text{otherwise.} \end{cases}$$

Remark: Consequently the Chebyshev constant is

$$C_\alpha = \frac{2\,\overline{F}(z_\alpha) z_\alpha^\alpha}{M_\alpha}.$$

If f is differentiable on $(0, +\infty)$, and $\frac{zf'(z)}{f(z)}$ is a strictly decreasing function of $z > 0$, then our condition is satisfied, and the positive root of equation defining z_α is unique.

Some particular choices of base distributions, for which the conditions are satisfied:
$U[-1, 1]$, standard normal, Student's t-distribution with $\nu > 0$ degrees of freedom, two-sided gamma distribution