COWLES COMMISSION DISCUSSION PAPER: STATISTICS NO. 362

Note on Generalized Convex Functions and the Decision Problem

Roy Radner

July 17, 1951

1. Introduction.

The purpose of this note is to give the proofs of two theorems used in Cowles Commission Discussion Paper: Economics No. 2018 (Theorems 1 and 2, p. 2.11). The theorems proved here are actually more general than those quoted in that paper; they are stated and proved in purely function theoretic form in paragraphs 2 and 3, and translated into the language of the theory of decision-making in paragraph 4. The concepts mentioned in this last paragraph are more fully discussed in Economics 2018, section 1.B.

The use of the term "quasi-convex functions" in Economics 2018 was unfortunate, since other authors (e.g., Slater, Debreu) have used the term in different contexts. In this note such functions are merely called "convex."

Theorem 1 below is a geometrically obvious extension of a well known result on convex functions, but the analytic proof given here is as long and tedious as it is elementary; nevertheless, the author has felt it desirable to include an analytic proof for the sake of completeness. Suggestions leading to a simple, direct, yet reasonably rigorous proof would be appreciated.

2. Throughout the rest of this note we will consider a fixed closed subset K of the real numbers $\mathring{E}$. We will also denote the closed unit interval [0,1] by I.

First some definitions: [1]

Def. 2.1. A real valued function u defined on K is convex on K if

$$\left.\begin{array}{l} k_1, k_2 \text{ in } K \\ a \text{ in } I \\ a\,k_1 + (1-a)\,k_2 \text{ in } K \end{array}\right\} \Longrightarrow \left\{\begin{array}{l} u\,(ak_1 + (1-a)\,k_2) \leq \\ \quad au\,(k_1) + (1-a)\,u\,(k_2) \end{array}\right.$$

Def. 2.2. (h,k) is a simple pair of points in K if $h \leq k$ and there exists no ℓ in K such that $h < \ell < k$.

Def. 2.3. f is a simple (probability) distribution on K if f assigns probability one to some simple pair in K.

Def. 2.4. For any x in E $\left\{\begin{array}{l} \underline{x} = \ell ub \ \{k \mid k \in K,\ k \leq x\} \\ \bar{x} = g\ell b \ \{k \mid k \in K,\ k \geq x\} \end{array}\right.$

K closed implies that $\underline{x}$ and $\bar{x}$ are in K.

Def. 2.5. A real number x is bounded by K if there exist k_1 and k_2 in K such that $k_1 \leq x \leq k_2$.

Theorem 1.

If u convex on K, f a distribution on K with mean μ bounded by K; then there exists a simple distribution g on K with mean μ, and

$$E_g\,(u(k)) \leq E_f\,(u(k))$$

We prove the theorem in a series of lemmas.

Lemma 2.1. If a convex on K, $h_1 \geq h_2 \geq k$ in K, a_1, a_2 in I, and $a_1 h_1 + (1-a_1)\,k = a_2 h_2 + (1-a_2)\,k$;

Then $a_2\,u\,(h_2) + (1-a_2)\,u\,(k) \leq a_1\,u\,(h_1) + (1-a_1)\,u\,(k)$

---

1. The meanings of the mathematical symbols used are given at the end of the paper.

Proof: There exists b in I such that $b h_1 + (1-b) k = h_2$.

By convexity of u, $u(h_2) \leqq b\, u(h_1) + (1-b)\, u(k)$

Then $a_2 u(h_2) + (1-a_2) u(k) \leq a_2 (b\, u(h_1) + (1-b)\, u(k)) + (1-a)\, u(k)$

$$\leq a_2 \; b\, u(h_1) + [a_2 (1-b) + (1-a_2)] \; u(k)$$

But $a_2 b h_1 + [a_2 (1-b) + (1-a_2)] k = a_2 (b h_1 + (1-b) k) + (1-a_2) k$

$$= a_2 h_2 + (1-a_2) k$$

$$= a_1 h_1 + (1-a_1) k$$

Then $a_2 b = a_1$ and $a_2 (1-b) + (1-a) = (1-a_1)$. QED.

Lemma 2.2.

$$\left.\begin{array}{l} u \text{ convex on } K \\ h_1 \leq h_2 \leq k \text{ in } K \\ x = a h_1 + (1-a) h_2 \\ y = b x + (1-b) k \\ y = c h_2 + (1-c) k \\ a, b, c \text{ in } I \end{array}\right\} \Longrightarrow \left\{\begin{array}{l} v \stackrel{=}{} b\,[a\, u(h_1) + (1-a)\, u(h_2)] + (1-b)\, u(k) \\ \qquad \geq c\, u(h_2) + (1-c)\, u(k) \end{array}\right.$$

Proof: There exists d in I such that $d h_1 + (1-d) k = y$. By Lemma 2.1,

(2.1) $$c\, u(h_2) + (1-c)\, u(k) \leq d\, u(h_1) + (1-d)\, u(k)$$

Let p be defined by: $v = p\,[d\, u(h_1) + (1-d)\, u(k)] + (1-p)\,[c\, u(h_2) + (1-c)\, u(k)]$.

It can be verified that $p = \dfrac{a c}{a c + (1-a) d}$.

Hence $0 \leq p \leq 1$ and the lemma follows from equation (2.1).

Lemma 2.3. If u convex on K, and x in E, $x = a \underline{x} + (1-a) \bar{\bar{x}}$, (a in I), define $u(x) \stackrel{=}{} a\, u(\underline{x}) + (1-a)\, u(\bar{\bar{x}})$. Then u is convex on the set obtained by adding x to K.

Proof: If $b x + (1-b) h = k \in K$, b in I, h in K, then apply Lemma 2.2 to get $u(k) \leq a\, u(x) + (1-a)\, u(h)$.

If $x = b\,h + (1-b)\,k$, $b \in I$, $h, k \in K$ then there exist c, d in I such that $\underline{x} = c\,h + (1-c)\,k$

$\bar{x} = d\,h + (1-d)\,k$

By convexity y u:

$$u(\underline{x}) \leq c\,u(h) + (1-c)\,u(k)$$

$$u(\bar{x}) \leq d\,u(h) + (1-d)\,u(k)$$

Substituting these inequalities in the definition of u (x) gives:

$$u(x) \leq [a\,c + (1-a)\,d]\,u(h) + [a\,(1-c) + (1-a)\,(1-d)]\,u(k)$$

It is easily verified that the coefficients of u (h) and u (k) in the above equation are equal to b and (1-b) respectively. QED.

Def. 2.6. If u is convex on K, we define u^*, an extension of u to E, by:

$$u^*(x) \equiv a_x\,u(\underline{x}) + 1-a_x)\,u(\bar{x}),$$

where a_x is defined by:

$$x = a_x\,\underline{x} + (1-a_x)\,\bar{x}$$

Lemma 2.4. If u convex on K, then u^* is convex on $[k_1, k_2]$ where $k_1 = g\ell b\ K$, $k_2 = \ell ub\ K$.[2]

Proof: By induction, using lemma 2.3, u^* is convex on any set obtained by adding a finite number of points to K. From the definition of convexity it is clear that a function is convex on an interval if and only if it convex on every set of three points in the interval. QED.

To complete the proof of the theorem, define the distribution g b. :

$$\Pr(k = \underline{\mu}) = a_\mu$$
$$\Pr(k = \bar{\mu}) = 1-a_\mu$$

Then by definition 2.6,

---

2. If $k_1 = -\infty$, then interpret $[k_1, k_2]$ to be $\{x \text{ in } E \mid -\infty < x \leq k_2\}$, etc.

$$u^*(E_f[x]) = u^*(\mu) = E_g[u(k)]$$

It is well known (Jensen's Inequality) that for a function u^* convex on $[k_1, k_2]$ and a distribution f on $[k_1, k_2]$

$$u^*(E_f[x]) \leq E_f[u^*(x)]$$ QED

3. We continue to consider a real-valued function u defined on K.

Def. 3.1. $\Delta_h u(k) = \dfrac{u(h) - u(k)}{h - k}$

Note that $\Delta_k u(h) = \Delta_h u(k)$

Lemma 3.1. u is convex on K if and only if for every k in K, $\Delta_h u(k)$ is a non-decreasing function of h in K.

Proof: Assume u is convex.

Case I, $k \leq h_1 \leq h_2$: There exists a in I such that

$$(3.1) \qquad \begin{aligned} h_1 &= a k + (1-a) h_2 \\ h_1 - k &= (a-1)(h_2 - k) \end{aligned}$$

By convexity, $u(h_1) \leq a\, u(k) + (1-a)\, u(h_2)$

$$u(h_1) - u(k) \leq (1-a)[u(h_2) - u(k)]$$

Dividing this by (3.1) gives

$$\Delta_{h_1} u(k) \leq \Delta_{h_2} u(k)$$

Case II, $h_1 \leq k \leq h_2$: } use same type of argument as in
Case III, $h_1 \leq h_2 \leq k$: } Case I.

The converse is proved by retracing the steps of the above proof. QED

Def. 3.2.

$$\Delta^+ u(k) = g\ell b \left\{\Delta_h u(k) \mid h > k, h \text{ in } K\right\}$$

$$\Delta^- u(k) = \ell ub \left\{\Delta_h u(k) \mid h < k, h \text{ in } K\right\}$$

Lemma 3.2. If u is convex on K, then both $\Delta^+ u(k)$ and $\Delta^- u(k)$ are non-decreasing functions of k in K, and $\Delta^- u(k) \leq \Delta^+ u(k)$ for every k in K.

<u>Proof</u>: To show $\Delta^+ u(k)$ non-decreasing, let $k < h$.

By lemma 3.1, for every $h_1 > h$:

$$\Delta^+ u(k) \leq \Delta_h u(k) = \Delta_k u(h) \leq \Delta_{h_1} u(h)$$

Hence $\quad \Delta^+ u(k) \leq \Delta^+ u(h)$.

The proof for $\Delta^- u(k)$ is similar.

That $\Delta^- u(k) \leq \Delta^+ u(k)$ follows immediately from lemma 3.1.

<u>Lemma 3.3</u>. If u is convex on K and is minimum at K_o, then $u(k)$ is non-increasing in k for $k < k_o$ and is non-decreasing for $k > k_o$.

Proof immediate from lemma 3.1.

Consider now a real-valued function $u(k, p)$ where k in K, p in P and P is any non-empty set. We will regard p as a parameter, and it will be understood that the notion of convexity and the operations Δ_h, Δ^+, Δ^-, are to be applied to $u(k, p)$ as a function of k for fixed p. From the definitions we immediately get the following:

<u>Lemma 3.4</u>. If $u(k, p)$ is convex on K for every p in P, then $\sup_p u(k, p)$ is convex on K.

<u>Theorem 2</u>. If $u(k, p)$ is convex on K for every p in P, $u^*(x, p)$ is defined as in Def. 2.6, and $\sup_p u^*(x, p)$ is minimum at x_o, then:

1) $\min_x \sup_p u^*(x, p) = \sup_p [a_{x_o} u(\underline{x}_o, p) + (1-a_{x_o}) u(\bar{x}_o, p)]$ where $\underline{x}_o$, $\bar{x}_o$, a_{x_o} are as defined in Definitions 2.4 and 2.6.

2) $\sup_p u(k, p)$ is minimum with respect to k in K at either $\underline{x}_o$ or $\bar{x}_o$.

3) If, further, $\max_p u(k, p)$ exists for every k in K, then

$$\bar{x}_o = g\ell b \left\{ k \text{ in } K \mid \Delta^+ u(k, p_k) > 0 \right\}$$

where p_k is defined by $u(k, p_k) = \max_p u(k, p)$

Proof: 1) This is true by the definition of u^*.

2) This follows from lemmas 3.3 and 3.4.

3) First we show that if $k \geq h$, then $\Delta^+ u\,(k, p_k) \geq \Delta^+ u\,(h, p_h)$.

For every $\ell \geq h$, by lemma 3.1:

$$u_\ell(k, p_k) \geq \Delta_h\, u\,(k, p_k)$$

Hence $\Delta^+ u\,(k, p_k) \geq \Delta_h\, u\,(k, p_k) = \dfrac{u\,(h, p_k) - u\,(k, p_k)}{h - k} \geq$

$$\geq \frac{u\,(h, p_h) - u\,(k, p_h)}{h - k} = \Delta_k\, u\,(h, p_h) \geq \Delta^+ u\,(h, p_h).$$

Let $k_o = g\ell b \left\{k \text{ in } K \mid \Delta^+ u\,(k, p_k) > 0\right\}$. Then for every $k > k_o$, $\Delta^+ u\,(k, p_k) > 0$.

If k_o is a limit point of $\left\{k \text{ in } K \mid k > k_o\right\}$ then there is no simple pair (k_o, k) such that $k > k_o$. If k_o is not a limit point of $\left\{k \text{ in } K \mid k > k_o\right\}$ then by definition of k_o, $\Delta^+ u\,(k_o, p_{k_o}) > 0$.

We will show that for every simple pair (h_1, h_2) in K such that $k_o \leq h_1 < h_2$:

$$\min_{h_1 \leq x \leq h_2}\ \max_p\ u^*(x, p) = u\,(h_1, p_{h_1}).$$

Since $h_1 \geq k_o$, $\Delta^+ u\,(h_1, p_{h_1}) > 0$, and $u(h_2, p_{h_1}) > u(h_1, p_{h_1})$.

Hence for any $x = a\,h_1 + (1-a)\,h_2$, a in I;

$$\begin{aligned}\max_p u^*(x, p) &= \max_p\,[a\,u\,(h_1, p) + (1-a)\,u\,(h_2, p)]\\ &\geq a\,u\,(h_1, p_{h_1}) + (1-a)\,u\,(h_2, p_{h_1})\\ &> u\,(h_1, p_{h_1})\end{aligned}$$

Furthermore, since for $k > k_o$, $\Delta^+ u\,(k, p_k) > 0$, for $k > k_o$ $u(k, p_k) > u(k_o, p_{k_o})$.

Hence we have shown that

$$\min_{x \geq k_o} \max_{p} u^* (x, p) = u (k_o, p_{k_o}).$$

In a similar fashion one proves that

$$\min_{x \leq k_1} \max_{p} u^* (x, p) = u (k_1, p_{k_o}),$$

where $k_1 = \ell ub \{ k \text{ in } K \mid k < k_o \}$. Hence $k_o = \bar{x}_o$ and $k_1 = \underline{x}_o$. QED.

3. Applications.

Let $U(k, p)$ be the expected utility to the decision maker associated with the strategy k (in K) and the state of nature p (in P), where K is a (non-empty) closed subset of the real numbers and P is any (non-empty) set. If we let K^* be the set of all probability distributions on K, then K^* represents the set of all mixed strategies obtained from the set K of pure strategies. For any f in K^*, the expected utility to the decision maker associated with f and any p in P is $E_f [U (k, p)]$, the expected value of $U (k, p)$ relative to f.

Def. 3.1. A subset L^* of K^* is complete in K^* if for every g in K^* there is an f in L^* such that $E_g [U (k, p)] \leq E_f [U (k, p)]$ for every p in P.

It is obvious that if H^* is complete in L^*, and L^* is complete in K^*, then H^* is complete in K^*.

The following is an immediate corollary of theorem 1.

Theorem 1'. If $-U (k, p)$ is convex on K for every p in P, and the set of all distributions in K^* with mean bounded by K is complete in K^*, then the set of all simple distributions in K^* is complete in K^*.

The regret function associated with $U (k, p)$ is defined for f in K by:

$$R^* (f, p) = \sup_{g \in K^*} E_g [U (k, p)] - E_f [U (k, p)]$$

Def. 3.2. An element $\hat{f}$ in K^* is called a minimax regret strategy relative to K^* if:

$$\sup_p R(\hat{f}, p) \leq \sup_p R(f, p) \quad \text{for all } f \text{ in } K^*.$$

It is clear that if L^* is complete in K^*, and there exists a minimax regret strategy relative to K^*, then there exists an element of L^* which is a minimax regret strategy relative to K^*.

For every k in K, K^* contains a distribution which assigns probability one to K, hence:

$$\sup_{f \in K^*} E_f [U(k, p)] = \sup_{k \in K} U(k, p)$$

and $R(f, p) = \sup_{h \in K} U(h, p) - E_f [U(k, p)]$

$$R(f, p) = E_f [\sup_{h \in K} U(h, p) - U(k, p)]$$

It is suggestive to denote the expression in the brackets in the above equation by R (k, p) (although purists will object). This would be the regret function if mixed strategies were not available. Thus:

$$R(f, p) = E_f [R(k, p)] \tag{4.1}$$

It is clear that, for every p, R (k, p) is convex on K if and only if $- U(k, p)$ is convex on K. Let R (k, p) be convex on K for every p, let $R^*(k, p)$ be defined as in Def. 2.6, and let g be a simple function on K with mean m. Then from (4.1)

$$R(g, p) = R^*(m, p)$$

Thus we have the following corollary of Theorem 2:

Theorem 2'. If R (k, p) is convex on K for every p in P, the set of all distributions in K^* which have mean bounded by K is complete in K^*, and there exists a minimax regret strategy relative to K^*, then

1) $\min\limits_{f \in K^*} \sup\limits_{p \in P} R(f, p) = \sup\limits_{p} [a_{x_o} R(\underline{x}_o, p) + (1-a_{x_o}) R(\bar{x}_o, p)]$

where $\sup\limits_{p} R^*(x, p)$ is minimum at x_o, and $\underline{x}_o$, $\bar{x}_o$, a_{x_o} are defined as in Definitions 2.4 and 2.6.

2) $\sup\limits_{p} R(k, p)$ is minimum with respect to k in K at either $\underline{x}_o$ or $\bar{x}_o$.

3) If, further, $\max\limits_{p} R(k, p)$ exists for every k in K, then:

$$\bar{x}_o = g\ell b \left\{ k \text{ in } K \mid \Delta^+ R(k, p_k) > 0 \right\},$$ where p_k is defined by $R(k, p_k) = \max\limits_{p} R(k, p)$.

Glossary of Symbols

$\in$	:	in
$\Rightarrow$	:	implies
$\equiv$	:	equals by definition
ℓub	:	least upper bound
$g\ell b$	:	greatest lower bound
$\{x \mid A.B.C\}$	:	the set of all x such that A, B, C hold.
$E_f[u]$	:	the expected value of u relative to the probability distribution f.