A GENERAL MEAN VALUE THEOREM

BY E. D. CASHWELL AND C. J. EVERETT

Communicated by Jürgen K. Moser, May 17, 1968

We present here in general terms the idea of the mean of a function relative to a "weight function" \( w(\xi, \nu) \), special instances and applications appearing elsewhere [1], [2].

1. The weight function. If \( X = [h, k] \) is a real interval, \( (I, A, \mu) \) a finite measure space with \( \mu(I) = 1 \), and \( w(\xi, \nu) \) a nonnegative function on \( X \times I \) which, for each \( \nu \) of \( I \), is measurable, and positive a.e. on \( X \), then the indefinite integral

\[
W(x, \nu) = \int_h^x w(\xi, \nu) d\xi
\]

is defined on \( X \times I \), and the function

\[
\mathcal{W}(x) = \int_I W(x, \nu) d\mu, \quad x \in X
\]

which we assume to exist, is continuous and strictly increasing on \( X \), as is \( W(x, \nu) \) for each \( \nu \).

2. The mean of a function. Let \( x(\nu) \) be any \( \mu \)-integrable function on \( I \) to \( X \) for which the integral functional

\[
\mathcal{W}_\nu = \int_I W(x(\nu), \nu) d\mu
\]

exists. Let \( x_u \) be the essential upper bound of \( x(\nu) \) on \( I \), i.e., the g.l.b. of all real \( x \) for which \( \mu \{ \nu | x(\nu) > x \} = 0 \), the essential lower bound

\[\text{1 Work performed under the auspices of the U. S. Atomic Energy Commission.}\]
\( x_i \) being analogously defined. Clearly \( x(\nu) \) is constant \( \mu \text{-a.e.} \) if and only if \( x_i = x_u \).

Referring to (1), it is apparent that the continuous, strictly increasing function

\[
B(x) = \int_I \int_{z(\nu)}^x w(\xi, \nu) d\xi d\mu = \nu(x) - \nu_z
\]

has a unique zero \( b \) on \( X \), namely

\[
b = \nu^{-1}(\nu_z)
\]

called the mean of \( x(\nu) \) relative to \( w(\xi, \nu) \). For, if \( x(\nu) \) is a constant \( x_0, \mu \text{-a.e.} \), we have \( B(x_0) = 0 \); otherwise we see that \( B(x_i) < 0 < B(x_u) \), so that \( B(b) = 0 \) for a unique \( b \) on \( (x_i, x_u) \).

3. The principal theorem. For an arbitrary bounded, monotone nondecreasing function \( g(\xi) \) on \( X \), we analogously define

\[
G(x, \nu) = \int_I \int_{z(\nu)}^x g(\xi) w(\xi, \nu) d\xi
\]
on \( X \times I \), and assume the existence of

\[
\mathcal{G}(x) = \int_I G(x, \nu) d\mu, \quad x \in X
\]
and of

\[
\mathcal{G}_z = \int_I G(x(\nu), \nu) d\mu.
\]
For the function

\[
C(x) = \int_I \int_{z(\nu)}^x g(\xi) w(\xi, \nu) d\xi d\mu = \mathcal{G}(x) - \mathcal{G}_z
\]
we then have the basic

**Theorem.**

(2) \( C(b) \leq 0 \)
or

\[
\mathcal{G}(\nu^{-1}(\nu_z)) \leq \mathcal{G}_z.
\]

Equality holds if and only if \( x(\nu) = b, \mu \text{-a.e.} \), or \( g(\xi) = g(b) \) everywhere on the open interval \( (x_i, x_u) \).
The inequality is rendered transparent by splitting \( I \) into the \( \mu \)-measurable subsets \( L, Z, U \) on which \( x(\nu) \leq b \), respectively, and observing that

\[
-C(b) = g(b)B(b) - C(b) \\
= \int_L \int_{x(\nu)}^{b} \{g(b) - g(\xi)\} w(\xi, \nu) d \xi d \mu \\
+ \int_U \int_{b}^{x(\nu)} \{g(\xi) - g(b)\} w(\xi, \nu) d \xi d \mu \geq 0.
\]

4. Two applications. In the simplest case, \( w(\xi, \nu) = 1 \), (2) is Jensen's inequality

\[
G\left( \int_I x(\nu) d \mu \right) \leq \int_I G(x(\nu)) d \mu
\]

for the general convex function \( G(x) = G(x, \nu) = \int_h^b g(\xi) d \xi \) [3, §13.34, §18.43]. A particular instance is mentioned in [2, §3].

Again, if we take \( h > 0 \) and set \( w(\xi, \nu) = \xi^{-1} \), \( \nu \) real, we find that \( b \) is the "mean of order \( s \)" of \( x(\nu) \):

\[
M_s = \left\{ (\int_I x^s(\nu) d \mu)^{1/s} \right\}, \; s \neq 0; \exp \int_I \log x(\nu) d \mu, \; s = 0
\]

and (2) yields the classical inequality

\[
M_s \leq M_t \quad \text{for } s < t
\]

if one takes \( g(\xi) = \xi^{-s} \) [1], [2].

These are trivial examples of the "separable" case \( w(\xi, \nu) = w(\xi)f(\nu) \). Nonseparable cases arise naturally in physical problems, as indicated below.

5. A "minimax" principle. Let \( y(\nu) \) be a second function such as \( x(\nu) \), and suppose that

\[
\int_I \int_{x(\nu)}^{y(\nu)} w(\xi, \nu) d \xi d \mu = 0.
\]

This is equivalent to the assertion that \( y(\nu) \) and \( x(\nu) \) have the same mean \( b \) relative to \( w(\xi, \nu) \), and it follows at once from the Theorem (applied to \( y(\nu) \)) that
If we regard \( x(\nu) \) as initial temperature distribution on an interval \( I \), of mass \( m(\nu) \) per unit length \( (d\mu = m(\nu) d\nu) \), and specific heat \( w(\xi, \nu) \), then (3) singles out the energy conserving distributions \( y(\nu) \), and (4) (with \( g(\xi) = -1/\xi \)) shows that, among these, the entropy change is greatest for the uniform mean temperature \( y(\nu) \equiv \bar{b} \).

REFERENCES


2. ———, The mean of a function \( x(\nu) \) relative to a function \( w(\xi, \nu) \), Amer. Math. Monthly (to appear).