Principal Value Mean

Mark Elmer, Magdalena Mosbo, Scott Preston
SUNY Oswego

Abstract

Principal value mean (PVM) of a probability distribution is the mean based on Cauchy principal value integral. We show that principal value mean often behaves well under translation, and that this property guarantees that PVM is essentially unique in certain sense. We also construct a distribution whose PVM is not translation invariant. Finally, we discuss some problems with statistical relevance of PVM.

1 Introduction

The Cauchy distribution has density function

$$f(x) = \frac{1}{\pi(1+x^2)}.$$

The mean of Cauchy distribution does not exist, because the integral

$$\mu = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{x}{1+x^2}dx$$

diverges. However, the distribution is symmetric, so clearly 0 is a nice ”balance point”.One way to get around the convergence problem is to use the Cauchy principal value integral, defined (for example, in Kaplan, 2003) by

$$PV\int_{-\infty}^{\infty} f = \lim_{N\to\infty}\int_{-N}^{N} f(x)dx$$

Using this concept, we may define the *principal value mean* of a distribution f as

$$\mu_{PV} = \lim_{N\to\infty}\int_{-N}^{N} xf(x)dx \tag{1}$$

if the limit exists. Now certainly if a mean of distribution exists, then so does a principal value mean and they are equal.

Some problems arise in connection to this definition. One is the uniqueness of such concept. The function is integrated over expanding intervals $[-N.N]$ which are centered at 0. The choice of 0 might seem arbitrary. So a natural question is: would $\lim_{N\to\infty}\int_{m-N}^{m+N} xf(x)dx$ have the same value for every m? We will show that the ”uniqueness” of PVM is equivalent to translation invariance, and that most reasonable distributions have these properties. This is good news - it means that PVM could be considered a reasonable center of distribution!

In Section 3, we construct a distribution whose principal value mean is not translation invariant. For such distribution,

$$\lim_{N\to\infty}\int_{m-N}^{m+N} xf(x)dx \tag{2}$$

depends on m. However, we also show that the limit cannot assume two different finite values; that is, in the worst case, it can be finite for some m and not exist for others. So, in this generalized sense, PVM is unique.

2 Translation invariance

Suppose f is a density function for a distribution having principal value mean μ_{PV}. Let f_m denote the translation of f: $f_m(y) = f(y-m)$. If for any real number m the principal value mean for f_m is $\mu_{PV} + m$, we say that f has a *translation invariant principal value mean.*

Clearly, if a distribution has a finite mean μ, then its principal value mean is equal to μ and is translation invariant. It can also be shown (via straightforward, but not exactly easy calculation) that the Cauchy distribution has a translation invariant principal value mean. The natural questions then is whether every distribution have a translation invariant principal value mean.

The first results shows that "uniqueness" (in the sense of Section 1) and translation invariance are equivalent. We also prove that these properties are equivalent to another condition, which we will use in future proofs.

Theorem 1: Let f be a distribution with a principal value mean $\mu = \mu_{PV}$ as defined in (1). Then for every real number m, the following conditions are equivalent:

1. $\lim_{N\to\infty}\int_{-N}^{N} xf(x-m)dx = \mu + m$.
2. $\lim_{N\to\infty}\int_{-m-N}^{-m+N} xf(x)dx = \mu$.
3. $\lim_{N\to\infty}\int_{N-m}^{N+m} xf(x)dx = 0$

Proof: To prove the equivalence of the first two conditions, observe that

$$\int_{-N}^{N} xf(x-m)dx = \int_{-m-N}^{-m+N}(u+m)f(u)dx = \int_{-m-N}^{-m+N} uf(u)dx + m\int_{-m-N}^{-m+N} f(u)dx$$

Since $\int_{-\infty}^{\infty} f = 1$, we know that $\lim_{N\to\infty}\int_{-m-N}^{-m+N} f(u)dx = 1$ so

$$\lim_{N\to\infty}\int_{-N}^{N} xf(x-m)dx = \lim_{N\to\infty}\int_{-m-N}^{-m+N} xf(x)dx + m$$

which shows that the conditions 1. and 2. are equivalent.

Now to prove the equivalence of 2. and 3., note that

$$\int_{m-N}^{m+N} xf(x)dx = \int_{m-N}^{N-m} xf(x)dx + \int_{N-m}^{m+N} xf(x)dx$$

Since the intervals $[m-N, N-m]$ are symmetric, we know

$$\lim_{N\to\infty} \int_{m-N}^{N-m} xf(x)dx = \mu$$

Therefore,

$$\lim_{N\to\infty} \int_{m-N}^{m+N} xf(x)dx = \mu \quad \text{iff} \quad \lim_{N\to\infty} \int_{N-m}^{N+m} xf(x)dx = 0$$

In the next theorem we will show that for most reasonable distributions (Cauchy distribution being one of them), the pricipal value mean is translation invariant.

Theorem 2: Let f be a distribution function such that

$$\lim_{x\to\infty} xf(x) = 0 \quad \text{or} \quad \lim_{x\to-\infty} xf(x) = 0$$

Then f has translation invariant principal value mean.

Proof: Observe that $f(x)$ has translation invariant principal value mean if and only if $f(-x)$ does, so the second asserion will follow from the first. Now suppose that $\lim_{x\to\infty} xf(x) = 0$ and choose real numbers m and ϵ, with $\epsilon > 0$. We may find N large enough so that $N - |m| > 0$ and $0 < xf(x) < \frac{\epsilon}{2|m|}$ for all x in the interval $[N-|m|, N+|m|]$. Therefore

$$|\int_{N-m}^{N+m} xf(x)dx| \leq \frac{\epsilon}{2|m|}|(N+m)-(N-m)| = \epsilon$$

We conclude that

$$\lim_{N\to\infty} \int_{N-m}^{N+m} xf(x)dx = 0$$

so, by Theorem 1, f has translation invariant principal value mean.

Most reasonable density functions, including that for Cauchy distribution, satisfy the hypothesis of Theorem 2, so translation invariance for principal value mean is typical.

3 Principal value means without translation invariance

Lemma 1: Let $(a_n)_{n=1}^{\infty}$ be a sequence of real numbers such that $a_n \geq 0$ and $\sum_{n=1}^{\infty} a_n$ converges. Then, if $\lim_{n\to\infty} na_n$ exists, it must be 0.

Proof: This follows directly from limit comparison test. If the limit exists and is not 0, then by comparison with the harmonic series, the series $\sum_{n=0}^{\infty} a_n$ would diverge.

Theorem 3: Let f be a distribution function and let m be a real number. Then the limit

$$\lim_{N\to\infty} \int_{N-m}^{N+m} xf(x)dx$$

is either 0 or does not exist.

Proof: First assume $m > 0$. Define the sequences

$$I_N = \int_{N-m}^{N+m} xf(x)dx \quad \text{and} \quad a_N = \int_{N-m}^{N+m} f(x)dx$$

Notice that since $f(x) \geq 0$, for any positive x we have also $xf(x) \geq 0$. So, for N large enough, we have

$$(N-m)a_N \leq I_N \leq (N+m)a_N$$

and so

$$-ma_N \leq I_N - Na_N \leq ma_N$$

We will now show that $\sum_{N=0}^{\infty} a_N$ converges. To this end, define

$$f_N(x) = f(x)\chi_{[N-m,N+m]}(x)$$

Then for every fixed x, $f_N(x)$ is non-zero only when $x - m \leq N \leq x + m$, that is, for no more than $2m + 1$ values of N. So, for each x,

$$\sum_{N=0}^{\infty} f_N(x) \leq (2m+1)f(x)$$

Hence,

$$\sum_{N=0}^{\infty} a_N = \sum_{N=0}^{\infty} \int_{N-m}^{N+m} f(x)dx = \int_{-\infty}^{\infty} \sum_{N=0}^{\infty} f_N(x) \leq (2m+1) \int_{-\infty}^{\infty} f(x)dx < \infty$$

Since the series $\sum_{N=0}^{\infty} a_N$ converges, we know $\lim_{N\to\infty} a_N = 0$. By the squeeze law,

$$\lim_{N\to\infty} I_N - Na_N = 0$$

Thus, the limits $\lim_{N\to\infty} I_N$ and $\lim_{N\to\infty} Na_N$ are either equal, or neither limit exists. By Lemma 1, $\lim_{N\to\infty} Na_N$ is either 0 or does not exist, so the same applies to $\lim_{N\to\infty} I_N$.

Now assume that $m < 0$. The proof then applies to $-m$, and hence

$$\lim_{N\to\infty} \int_{N+m}^{N-m} xf(x)dx$$

is either 0 or undefined. However, since

$$\int_{N-m}^{N+m} xf(x)dx = -\int_{N+m}^{N-m} xf(x)dx$$

so the result holds for $m < 0$ as well.

Corollary 2: Let f be a distribution such that $\mu = \mu_{PV}$ exists. Let m be a real number. Then the limit

$$\lim_{N\to\infty} \int_{m-N}^{m+N} xf(x)dx$$

is either equal to μ or is undefined. In other words, a distribution cannot have two finite, different principal value means.

It is possible, however, to find a distribution whose principal value mean exists, but is not translation invariant. The next example presents such a distribution.

Example 1: Define f as follows:

$$f(x) = \frac{c}{k^2} \quad \text{if} \quad |x| \in [k^2 - \frac{1}{2}, k^2 + \frac{1}{2}]$$

and let $f(x) = 0$ otherwise. Note that

$$\int_{-\infty}^{\infty} f(t)dt = 2c \sum_{k=1}^{\infty} \frac{1}{k^2}$$

so the constant c can be found so that f is a density function. Also, since f is even, its principal value mean must be 0. However, it can be shown that

$$\lim_{n\to\infty} \int_{n-\frac{1}{2}}^{n+\frac{1}{2}} xf(x)$$

does not exist. Indeed, if $n > 1$ is not a perfect square, then $f(x) = 0$ on the interval $[n - \frac{1}{2}, n + \frac{1}{2}]$ so $I_n = 0$. However, if $n = k^2$, then

$$I_n = \int_{n-\frac{1}{2}}^{n+\frac{1}{2}} xf(x) = \int_{n-\frac{1}{2}}^{n+\frac{1}{2}} x\frac{c}{k^2}dx \geq (k^2 - \frac{1}{2}) \int_{n-\frac{1}{2}}^{n+\frac{1}{2}} \frac{c}{k^2}dx = c(1 - \frac{1}{2n})$$

So (I_n) has a subsequence converging to $C \neq 0$, as well as a subsequence converging to 0. Thus $\lim_{N\to\infty} I_N$ does not exist, and hence the principal value mean of f is not translation invariant.

4 Sampling from populations

A fundamental application of statistics involves characterization of the distribution of the sample mean $\bar{x}$ of n independent observations randomly drawn from a population with density f. The mean μ of a distribution is defined if and only if $\int_{-\infty}^{\infty} xf(x)dx$ converges absolutely. The strong law of large numbers establishes that the existence of the mean μ is necessary and sufficient for the convergence of $\bar{x}$ to μ with probability one.

Distributions f with no mean have a counterintuitive property: the sample mean does not, in all probability, get "closer and closer" to anything as the sample size increases. The Cauchy distribution is the archetype: the sample mean $\bar{x}$ actually has the same distribution as does a single observation X. (We have a rather involved proof of this using undergraduate mathematics; a proof of the result is almost trivial using characteristic functions (Billingsley, 1986).) This result is easily demonstrated with a simple simulation. The following graphs show typical results of such an experiment. Consider a sequence $(X_n)_{n=1}^{1000}$ of independent samples from a certain population, and let $\bar{x}_N = \frac{1}{N}\sum_{k=1}^{N} X_k$ be the average of the first N samples. The running averages from Normal distribution approach the mean; the running averages from Cauchy distribution are guaranteed not to approach anything. Any value in the sequence is outside the interval $[-1, 1]$ with probability 0.5.

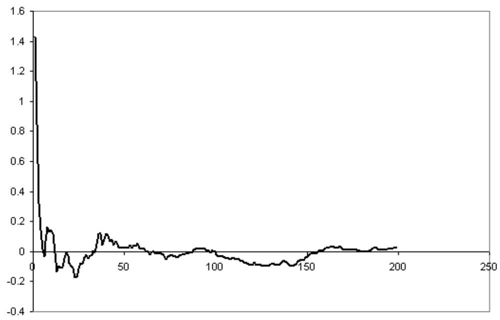

Fig. 1: Running averages from Normal distribution

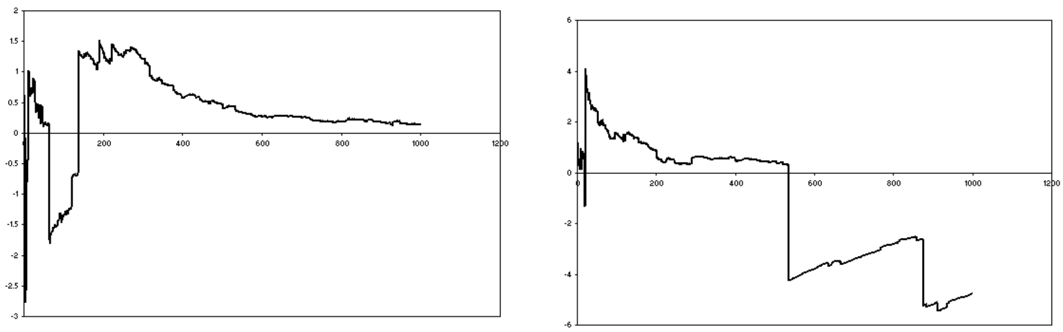

Fig. 2: Running averages from Cauchy distribution

Given the symmetry of the Cauchy distribution, this is curious at least. After all: each observation x occurs with density equal to that of $-x$ in this "population."

So, what happens in those cases where the mean μ does not exist? The key to an intuitive understanding the result is grounded in the analysis of integration (and hence summation). When $\int_{-\infty}^{\infty} x f(x)dx$ is not absolutely convergent, the limit of approximating sums fails to exist. The order of summation affects the limiting values. Given that the observations in our sample occur in a random order, we cannot ensure that the sample mean $\bar{x}$ converges to any particular value.

5 Conclusions

The concept of principal value mean appears to be a well-defined measure of the center of distribution. For most reasonable distributions, the principal value mean is translation invariant, which also means that the value of the limit (2) is independent of m. While there exist distributions whose principal value mean is not translation invariant, even for those distribtions (2) can only take on one finite value. This makes the principal value mean essentially well-defined concept worth further investigation. In particular, estimating a principal value mean should be explored.

References

[1] Billingsley, P.: Probability and Measure, 2nd edition, Wiley 1986

[2] Buck, C.: Advanced Calculus, etc.

[3] Kaplan, W.: Advanced Calculus, 5th edition, Addison-Wesley, 2003

[4] Harris, J. W. and Stocker, H. Handbook of Mathematics and Computational Science. New York: Springer-Verlag, 1998.