§1.2 Distribution Functions

WU Yuzhe

November 3, 2018

Problem 1. Let F be a distribution function (d.f.). Then for each x,

$$\lim_{\epsilon \downarrow 0} [F(x+\epsilon) - F(x-\epsilon)] = 0$$

unless x is a point of jump of F, in which case the limit is equal to the size of the jump.

Solution. We have $\lim_{\epsilon\downarrow 0} [F(x+\epsilon) - F(x-\epsilon)] = \lim_{\epsilon\downarrow 0} F(x+\epsilon) - \lim_{\epsilon\downarrow 0} F(x-\epsilon) = F(x+) - F(x-)$. If F is continuous at x, then $F(x+) = F(x-) \implies \lim_{\epsilon\downarrow 0} [F(x+\epsilon) - F(x-\epsilon)] = 0$. Otherwise F has a jump at x, and $\lim_{\epsilon\downarrow 0} [F(x+\epsilon) - F(x-\epsilon)] = F(x+) - F(x-)$ gives precisely the size of the jump. □

Problem 2. Let F be a d.f. with points of jump $\{a_j\}$. Prove that the sum

$$\sum_{x-\epsilon<a_j<x} [F(a_j) - F(a_j-)]$$

converges to zero as $\epsilon \downarrow 0$, for every x. What if the summation above is extended to $x - \epsilon < a_j \leq x$ instead? Give a proof of the continuity of F_c in Theorem 1 by using the problem.

Theorem 1. *Let F be a d.f. with points of jump $\{a_j\}$ and jump sizes $\{b_j\}$. Consider the function*

$$F_d(x) = \sum_j b_j \delta_{a_j}(x)$$

which represents the sum of all the jumps of F in the half-line $(-\infty, x]$, where the function δ_t is the point mass at t. Let

$$F_c(x) = F(x) - F_d(x);$$

then F_c is positive, increasing, and continuous.

Solution. We define $b_j = F(a_j) - F(a_j-)$ as in Theorem 1. Let $\eta > 0$ be given. Since the series $\sum_j b_j$ converges, there exists an integer N such that $\sum_{j=N+1}^{\infty} b_j < \eta$. Let $S = \{a_j\}_{j=1}^n \cap (-\infty, x)$. If $S = \varnothing$, then choose $\delta = 1$. Otherwise let $a^* = \sup S$ and choose $\delta = x - a^*$. In either case, we have if $0 < \epsilon < \delta$ and $x - \epsilon < a_j < x$, then $j > N$. Hence, $\sum_{x-\epsilon<a_j<x} b_j \leq \sum_{j=N+1}^{\infty} b_j < \eta$. We conclude that $\lim_{\epsilon\downarrow 0} \sum_{x-\epsilon<a_j<x} b_j = 0$.

If the summation above is extended to $x - \epsilon \leq x$, then the limit equals $F(x) - F(x-)$.

We now use the above result to prove the continuity of F_c in Theorem 1. Since both F and F_d are right continuous, so is F_c as their difference. Thus, we only need to establish the left continuity of F_c, or, equivalently, that $\lim_{\epsilon\downarrow 0}[F_c(x) - F_c(x-\epsilon)] = 0$. Note that

$$F_d(x) - F_d(x-\epsilon) = \sum_{x-\epsilon<a_j\leq x} b_j.$$

Therefore, we have

$$\begin{aligned}
&\lim_{\epsilon\downarrow 0}\,[F_c(x) - F_c(x-\epsilon)] \\
=&\lim_{\epsilon\downarrow 0}\,[F(x) - F(x-\epsilon)] - \lim_{\epsilon\downarrow 0}\,[F_d(x) - F_d(x-\epsilon)] \\
=&F(x) - F(x-) - \lim_{\epsilon\downarrow 0} \sum_{x-\epsilon<a_j\leq x} b_j \\
=&F(x) - F(x-) - (F(x) - F(x-)) = 0,
\end{aligned}$$

where the last equality comes from the result obtained in Problem 2. □

Problem 4. For a general increasing function F there is a similar decomposition $F = F_c + F_d$, where both F_c and F_d are increasing, F_c is continuous, and F_d is "purely jumping". [HINT: Let a be a point of continuity, put $F_d(a) = F(a)$, add jumps in (a, ∞) and subtract jumps in $(-\infty, a)$ to define F_d.]

Solution. Let a be a point of continuity and let $\{a_j\}$ be the countable collection of points of jump with jump sizes $\{b_j\}$. Define F_d as follows:

$$F_d(x) = \begin{cases} F(a) & \text{if } x = a, \\ F(a) + \sum_{a<a_j<x} b_j + F(x) - F(x-) & \text{if } x > a, \\ F(a) - \sum_{x<a_j<a} b_j + F(x) - F(x+) & \text{if } x < a, \end{cases}$$

where summation over an empty set is defined to be 0. Define $F_c = F - F_d$, We claim that F_c and F_d thus defined constitute a valid decomposition. Let $x < x'$, then we have

$$F_d(x') - F_d(x) = \sum_{x<a_j<x'} b_j + F(x') - F(x'-) + F(x+) - F(x) \leq F(x') - F(x).$$

It follows that both F_d and F_c are increasing. Let J be an open interval on which F is continuous, then there is no jump point in J. Let $y < y'$ be two points in J, then we have $F_d(y') - F_d(y) = \sum_{y<a_j<y'} b_j + F(y') - F(y'-) + F(y+) - F(y) = 0$. In other words, F_d is constant between jumps and hence is "purely jumping". We now show that F_c is continuous. Note that $F_d(x+) - F_d(x-) = \lim_{\epsilon\downarrow 0}[F_d(x+\epsilon) - F_d(x-\epsilon)] = F(x+) - F(x-)$ by a similar argument as elaborated in Problem 2. Therefore, we have $F_c(x+) - F_c(x-) = 0$ for all x, which completes the proof. □

Problem 6. A point x is said to belong to the support of the d.f. F iff for every $\epsilon > 0$ we have $F(x+\epsilon) - F(x-\epsilon) > 0$. The set of all such x is called the support of F. Show that each point of jump belongs to the support, and that each isolated point of the support is a point of jump. Give an example of a discrete d.f. whose support is the whole line.

Solution. Let x be a point of jump. For every $\epsilon > 0$, we have $F(x+) \leq F(x+\epsilon), F(x-) \geq F(x-\epsilon)$ and so $F(x+\epsilon) - F(x-\epsilon) \geq F(x+) - F(x-) > 0$. Hence x belongs to the support of F.

Let y be an isolated point of the support. Then there exists $d > 0$ such that every point in the open interval $(y, y+d)$ and in $(y-d, y)$ does not belong to the support of F. We claim that F is constant on $(y, y+d)$ and $(y-d, y)$. Suppose F is not constant on $(y, y+d)$. Then there exist $y < y_1 < y_2 < y+d$ such that $F(y_1) < F(y_2)$. Define $y^* := \inf A$, where $A = \{y' : F(y') > F(y_1)\}$. Clearly $y^* \in (y, y+d)$. For every $\epsilon > 0$, $y^* + \epsilon$ is not a lower bound of A. Therefore, there exists $y' < y^* + \epsilon$ such that $F(y') > F(y_1)$ and hence $F(y^* + \epsilon) \geq F(y') > F(y_1)$. On the other hand, $y^* - \epsilon \notin A$ since y^* is a lower bound of A. So $F(y^* - \epsilon) \leq F(y_1)$. Therefore, we have $F(y^* + \epsilon) - F(y^* - \epsilon) > 0$ and as a result $y^* \in (y, y+d)$ belongs to the support of F. But this contradicts the fact that no point of the support lies in the interval $(y, y+d)$. We conclude that F is constant on $(y, y+d)$. By a similar argument, we also have F is constant on $(y-d, y)$. It follows that $F(y+) = F(y+d/2)$ and $F(y-) = F(y-d/2)$. But $F(y+d/2) - F(y-d/2) > 0$ since y belongs to the support of F, and hence $F(y+) - F(y-) > 0$. We conclude that y is a point of jump.

We now give an example of a discrete d.f. whose support is the whole line. Let $\{q_j\}_{j=1}^{\infty}$ be an enumeration of the set of all rational numbers. Consider now

$$F(x) = \sum_{j=1}^{\infty} \frac{1}{2^j} \delta_{q_j}(x).$$

We have $\frac{1}{2^j} > 0$ for every j and $\sum_{j=1}^{\infty} \frac{1}{2^j} = 1$. Hence F is a discrete d.f. Let $x \in \mathbb{R}$ and let $\epsilon > 0$ be given. There exists a rational number $q_{j'} \in (x-\epsilon, x+\epsilon)$. We have $F(x+\epsilon) - F(x-\epsilon) = \sum_{x-\epsilon < q_j \leq x+\epsilon} \frac{1}{2^j} \geq \frac{1}{2^{j'}} > 0$. Hence x belongs to the support of F. We conclude that the support of F is the whole line. □

Problem 7. Prove that the support of any d.f. is a closed set, and the support of any continuous d.f. is a perfect set.

Solution. Let F be a d.f. and let A be its support. For every $x \in A^{\complement}$, there exists $\epsilon > 0$ such that $F(x+\epsilon) = F(x-\epsilon)$. Hence F is constant on $(x-\epsilon, x+\epsilon)$. It is clear that no point of support lies in $(x-\epsilon, x+\epsilon)$, i.e. $(x-\epsilon, x+\epsilon) \in A^{\complement}$. We conclude that $A^{\complement}$ is an open set, or, equivalently, that A is a closed set.

If F is now a continuous d.f., we already see that its support A is a closed set. To show that A is a perfect set, we need to show in addition that A contains no isolated point. But by Problem 6, each such point will be a point of jump, contradicting the fact that F is continuous. We conclude that the support of any continuous d.f. is a perfect set. □