Stat 8112 Lecture Notes

Big Oh Pee and Little Oh Pee

Charles J. Geyer

February 14, 2011

1 Big Oh and Little Oh

A sequence x_n of non-random vectors is said to be $O(1)$ if it is bounded and $o(1)$ if it converges to zero. If a_n is a sequence of non-random positive scalars, then

$$x_n = O(a_n) \tag{1a}$$

means

$$\frac{x_n}{a_n} = O(1)$$

(that is, x_n / a_n is bounded), and

$$x_n = o(a_n) \tag{1b}$$

means

$$\frac{x_n}{a_n} = o(1)$$

(that is, x_n / a_n converges to zero).

Note that the equals sign in (1a) and (1b) don't actually mean that anything is equal to anything. These "equations" are a convenient shorthand, but they are not about equality. Thus in careful arguments one needs to replace $O(a_n)$ and $o(a_n)$ by their definitions. Replace (1a) by there exists $M < \infty$ such that

$$\|x_n\| \le M a_n, \qquad \forall n \in \mathbb{N}.$$

Replace (1b) by

$$\frac{x_n}{a_n} \to 0, \qquad n \to \infty.$$

It is sometimes convenient to allow the sequences a_n to be vector-valued, in which case $O(a_n)$ means the same thing as $O(\|a_n\|)$ and $o(a_n)$ means the same thing as $o(\|a_n\|)$. That is, this is merely an abbreviation allowing us to omit the norm symbol.

We use the same notation to describe functions as well as sequences. We say that a function is $O(1)$ as $x \to 0$ if it is bounded on a neighborhood of zero, and we say it is $o(1)$ as $x \to 0$ if

$$f(x) \to 0, \qquad \text{as } x \to 0.$$

For any other function g

$$f(x) = O\bigl(g(x)\bigr)$$

means

$$\frac{f(x)}{\|g(x)\|} = O(1),$$

and

$$f(x) = o\bigl(g(x)\bigr)$$

means

$$\frac{f(x)}{\|g(x)\|} = o(1).$$

And we use similar notation for behavior at other points. For example, we say a Cauchy probability density function is $O(x^{-2})$ as $|x| \to \infty$.

2 Big Oh Pee and Little Oh Pee

A sequence X_n of random vectors is said to be $O_p(1)$ if it is bounded in probability (tight) and $o_p(1)$ if it converges in probability to zero.

The notations gain power when we consider pairs of sequences. Suppose X_n and Y_n are random sequences taking values in any normed vector space, then

$$X_n = O_p(Y_n) \tag{2a}$$

means $X_n / \|Y_n\|$ is bounded in probability and

$$X_n = o_p(Y_n) \tag{2b}$$

means $X_n / \|Y_n\|$ converges in probability to zero.

These notations are often used when the sequence Y_n is deterministic, for example, $X_n = O_p(n^{-1/2})$. They are also often used when the sequence Y_n is random, for example, we say two estimators $\hat{\theta}_n$ and $\tilde{\theta}_n$ of a parameter θ are *asymptotically equivalent* if

$$\hat{\theta}_n - \tilde{\theta}_n = o_p(\hat{\theta}_n - \theta)$$

and

$$\hat{\theta}_n - \tilde{\theta}_n = o_p(\tilde{\theta}_n - \theta).$$

Again the equals sign in (2a) and (2b) don't actually mean that anything is equal to anything. Thus in careful arguments one needs to replace (2a) by for every $\varepsilon > 0$ there exists an $M < \infty$ such that

$$\Pr(\|X_n\| \le M \|Y_n\|) \ge 1 - \varepsilon, \qquad \forall n \in \mathbb{N}$$

and replace (2b) by

$$\frac{X_n}{\|Y_n\|} \xrightarrow{P} 0, \qquad n \to \infty$$

or by for every $\varepsilon > 0$

$$\Pr(\|X_n\| \geq \varepsilon \|Y_n\|) \to 0.$$

We also use big oh and little oh and big oh pee and little oh pee notation for terms in equations (not just for right-hand sides of equations). For example, a function f is differentiable at x if

$$f(h + h) = f(x) + f'(x)h + o(h)$$

(see the handout on the delta method for more on this) and one case of Slutsky's theorem says

$$X_n \xrightarrow{w} X$$

implies

$$X_n + o_p(1) \xrightarrow{w} X.$$