STAT 830
Landau Notation: big O and little o

Richard Lockhart

Simon Fraser University

STAT 830 — Fall 2020

Purposes of These Notes

- Introduce Landau's formalism for "on the order of"
- Do O, o, O_P and o_P
- Present some of the algebraic rules

Notation for Size or ‘Order’ of sequences

Big “O”

Definition: If a_n and b_n are sequences of constants with $a_n > 0$ then

$$b_n = O(a_n)$$

means there is a constant M such that for all n

$$|b_n| \le M a_n.$$

Definition: If U_n is a sequence of random variables and $a_n > 0$ a sequence of constants then we write

$$U_n = O_P(a_n)$$

if, for each $\epsilon > 0$ there is an M (depending on ϵ but not n) such that

$$P(|U_n| > M|a_n|) < \epsilon$$

Notation for Size or 'Order' of sequences

Little "o"

If a_n and b_n are sequences of constants with $a_n > 0$ then
Definition: We say $b_n = o(a_n)$ if

$$\lim_{n\to\infty} \frac{b_n}{a_n} = 0.$$

Definition: We say $U_n = o_P(a_n)$ if $U_n/a_n \to 0$ in probability: for each $\epsilon > 0$

$$\lim_{n\to\infty} P(|U_n/a_n| > \epsilon) = 0.$$

Algebra Rules

- Can manipulate these algebraically if we are *careful*.

$$
\begin{aligned}
O(a_n)O(b_n) &= O(a_n b_n) \\
cO(a_n) &= O(a_n) \\
O(a_n)O_P(b_n) &= O_P(a_n b_n) \\
O_P(a_n)O_P(b_n) &= O_P(a_n b_n) \\
O(a_n) + O(b_n) &= O(\max\{a_n, b_n\}) \\
O_P(a_n) + O_P(b_n) &= O_P(\max\{a_n, b_n\})
\end{aligned}
$$

- All these rules hold with 'O' replaced by 'o'.
- In the rules both $a_n > 0$ and $b_n > 0$ and c is a constant not depending on n.

Blank Page for Algebra

Blank Page for Algebra

Big O with Little o

- The two go together sometimes:

$$
\begin{aligned}
O(a_n)o(b_n) &= o(a_n b_n) \\
o(O(a_n)) &= o(a_n) \\
o(a_n)O_P(b_n) &= o_P(a_n b_n) \\
o_P(a_n)O_P(b_n) &= o_P(a_n b_n) \\
o(a_n) + O(b_n) &= O(\max\{a_n, b_n\}) \\
o_P(a_n) + O_P(b_n) &= O_P(\max\{a_n, b_n\})
\end{aligned}
$$

- You can't cancel because each new occurence of O is different

$$O(a_n) - O(a_n) \neq 0.$$

- Only add and multiply and use positive rates.

Blank Page for Algebra

Blank Page for Algebra

Some examples and proof ideas

- Suppose Y_n is a sequence of random variables with mean 0 and variance a_n^2. Then
$$Y_n = O_P(a_n)$$
because
$$P\left(\frac{|Y_n|}{a_n} > M\right) \le \frac{a_n^2}{a_n^2 M^2} = \frac{1}{M^2}$$
which is less than ϵ if $M > 1/\sqrt{\epsilon}$.
- So in an iid sampling problem
$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} = O_P(1).$$

Blank Page for Algebra

More examples and proof ideas

- In the same problem s_n converges to σ in probability so

$$s_n = \sigma + o_P(1)$$

- Thus

$$\frac{\sigma}{s_n} = 1 + o_P(1)$$

- So

$$\begin{aligned}\frac{\sqrt{n}(\bar{X}_n - \mu)}{s_n} &= \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}\frac{\sigma}{s_n} \\ &= (1 + o_P(1))O_P(1) = O_P(1) + o_P(1) = O_P(1)\end{aligned}$$

Blank Page for Algebra

Blank Page for Algebra

More examples

- In a regular iid sampling problem each component of the score function is
$$O_P(\sqrt{n})$$
because the variance is n times the Fisher information in a single point
- The Hessian $H(\theta)$ is $O_P(n)$ and usually
$$H(\theta) = \mathcal{I}(\theta) + O_P(\sqrt{n})$$
and
$$H(\hat\theta) = \mathcal{I}(theta) + O_P(\sqrt{n}).$$
- The theory we proved said
$$\hat\theta = \theta_0 + O_P(1/\sqrt{n})$$
because
$$\hat\theta - \theta_0 = -H(\theta)^{-1}U(\theta) + o_P(1/\sqrt{n}).$$

Blank Page for Algebra

Blank Page for Algebra