Econometrics Problem Sets
William M. Boal

Problem Set 1
"Introduction: Subscripts and Summation Symbol"

(1.1) [Summation operator] Use the following dataset, consisting of three observations on x, to compute the formulas below.

Observation number (i)	x_i
1	8
2	3
3	7

a. Compute $\sum_{i=1}^{3} x_i$

b. Compute $\bar{x}$.

c. Compute $\sum_{i=1}^{3} x_i^2$.

d. Compute $\sum_{i=1}^{3} \left(x_i - \bar{x}\right)^2$.

(1.2) [Summation operator] Use the following dataset, consisting of three observations on x and y, to compute the formulas below.

Observation number (i)	x_i	y_i
1	3	2
2	4	5
3	8	5

a. Compute $\sum_{i=1}^{3} x_i$.

b. Compute $\bar{x}$.

c. Compute $\sum_{i=1}^{3} x_i^2$.

d. Compute $\sum_{i=1}^{3} \left(x_i - \bar{x}\right)^2$.

e. Compute $\sum_{i=1}^{3} x_i y_i$.

f. Compute $\sum_{i=1}^{3} \sum_{j=1}^{3} x_i y_j$.

(1.3) [Summation operator] Use the following dataset, consisting of four observations on x and y, to compute the formulas below.

Observation number (i)	x_i	y_i
1	2	3
2	5	1
3	3	2
4	6	5

a. Compute $\sum_{i=1}^{4} x_i$.

b. Compute $\bar{x}$.

c. Compute $\sum_{i=1}^{4} x_i^2$.

d. Compute $\sum_{i=1}^{4} \left(x_i - \bar{x}\right)^2$.

e. Compute $\sum_{i=1}^{4} x_i y_i$.

f. Compute $\sum_{i=1}^{4} \sum_{j=1}^{4} x_i y_j$.

(1.4) [Derivatives of sums] Let $f(\alpha) = \sum_{i=1}^{n} \frac{x_i}{\alpha}$, where the x_i are fixed but unspecified numbers (such as data) and α (the Greek letter alpha) is viewed as a variable. Find a formula for $df/d\alpha$, the derivative of this function with respect to α. [Hint: This will be a sum. Each term will be a formula in terms of α and x_i .]

(1.5) [Derivatives of sums] Let $f(\beta) = \sum_{i=1}^{n} \left(\beta + x_i y_i\right)^2$, where x_i and y_i are fixed but unspecified numbers (such as data) and β (the Greek letter beta) is viewed as a variable.

a. Find a formula for $df/d\beta$, the derivative of $f(\beta)$ with respect to β. [Hint: The answer is a sum. Each term will be a formula in terms of β and the x_i and y_i.]
b. Find a formula for the value of β (call it β^*) that minimizes $f(\beta)$. [Hint: The answer is a formula in terms of the x_i and y_i only.]

(1.6) [Derivatives of sums] Let $f(\gamma) = \sum_{i=1}^{n} \left(\gamma - x_i\right)^2$, where x_i are fixed but unspecified numbers (such as data) and γ (the Greek letter gamma) is viewed as a variable.

a. Find a formula for $df/d\gamma$, the derivative of $f(\gamma)$ with respect to γ. [Hint: The answer is a sum. Each term will be a formula in terms of γ and the x_i.]
b. Find a formula for the value of γ (call it γ^*) that minimizes $f(\gamma)$. [Hint: The answer is a formula in terms of the x_i only.]

(1.7) [Derivatives of sums] Let $f(\delta) = \sum_{i=1}^{n} (\delta - x_i^2)^2$, where x_i are fixed but unspecified numbers (such as data) and δ (the Greek letter delta) is viewed as a variable.

a. Find a formula for df/dδ , the derivative of f(δ) with respect to δ. [Hint: The answer is a sum. Each term will be a formula in terms of δ and the x_i.]
b. Find a formula for the value of δ (call it δ*) that minimizes f(δ). [Hint: The answer is a formula in terms of the x_i only.]

(1.8) [Derivatives of sums] Let $f(\delta) = -\frac{n}{2}\ln(\delta) - \sum_{i=1}^{n}\left(\frac{x_i^2}{2\delta}\right)$, where x_i are fixed but unspecified numbers (such as data) and δ (the Greek letter delta) is viewed as a variable The function ln(.) is the natural logarithm function—that is, the logarithm whose base is $e \approx 2.71828$.

a. Find a formula for df/dδ , the derivative of f(δ) with respect to δ. [Hint: The answer is a formula in terms of δ and the x_i.]
b. Find a formula for the value of δ (call it δ*) that minimizes f(δ). [Hint: The answer is a formula in terms of the x_i only.]

(1.9) [Summation identities] Suppose we have n observations on the variable w. Define as usual $\overline{w} = \frac{1}{n}\sum_{i=1}^{n} w_i$. Which of the following expressions are necessarily equal to zero?

a. $\sum_{i=1}^{n} w_i$.
b. $\sum_{i=1}^{n} w_i^2$.
c. $\left(\sum_{i=1}^{n} w_i\right) - n\overline{w}$.
d. $\sum_{i=1}^{n} (\overline{w} - w_i)$.
e. $\sum_{i=1}^{n} (-2(w_i - \overline{w}))$.
f. $\sum_{i=1}^{n} (\overline{w} - w_i)^2$.

(1.10) [Summation identities] Suppose we have n observations on the variable z. Define as usual $\overline{z} = \frac{1}{n}\sum_{i=1}^{n} z_i$. Which of the following expressions are necessarily equal to zero?

a. $\sum_{i=1}^{n} z_i$.
b. $\sum_{i=1}^{n} (z_i - \overline{z})$.
c. $\sum_{i=1}^{n} (z_i - \overline{z})^2$.
d. $\sum_{i=1}^{n} 2(z_i - \overline{z})$.
e. $n\overline{z} - \sum_{i=1}^{n} z_i$.
f. $\sum_{i=1}^{n} z_i^2$.

(1.11) [Summation identities] Suppose we have n observations on variables x and y. Define as usual

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad \text{and} \quad \overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \quad .$$

Prove the following, step by step.

$$\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})y_i$$

[Hint: Begin by using the distributive law, a (b+c) = ab + ac, to write:
$\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}) = \frac{1}{n}\sum_{i=1}^{n}\left[(x_i - \overline{x})y_i - (x_i - \overline{x})\overline{y}\right]$. Then break this up into two summations.]

(1.12) [Summation identities] Suppose we have n observations on variable w and m observations on variable z. Prove the following, step by step.

$$\sum_{i=1}^{n} \sum_{j=1}^{m}(w_i z_j) = \left(\sum_{i=1}^{n} w_i\right)\left(\sum_{j=1}^{m} z_j\right)$$

[Hint: Begin by noting that on the left side, w_i is a common factor in the summation over j , so we can write: $\sum_{i=1}^{n} \sum_{j=1}^{m}(w_i z_j) = \sum_{i=1}^{n}\left(w_i \sum_{j=1}^{m}(z_j)\right)$. Now what is the common factor of the summation over i ?]

[end of problem set]