Geometric Interpretation of Simpson’s Paradox

-Dhruv Desai

Problem

Prove that the inequalities $a_i/b_i < c_i/d_i$, $i=1,2,.....,n$ and $(\sum a_i)/(\sum b_i) > (\sum c_i)/(\sum d_i)$ can be satisfied simultaneously, provided at least one of c_i/d_i is less than $(\sum a_i)/(\sum b_i)$

Solution

-> The numbers can be represented as vectors **v1, v2, …….. Vn** with coordinates $(b_1,a_1),(b_2,a_2),.....,(b_n,a_n)$ and **v1',v2',......,vn'** with coordinates $(d_1,c_1),(d_2,c_2),...(d_n,c_n)$

->Hence, the problem becomes finding a set of vectors with angles $\theta_1,\theta_2,.....,\theta_n$ and $\theta_1',\theta_2',....\theta_n'$ such that resultant of unprimed vectors $\theta>\theta'$, provided at least one of $\theta_1',\theta_2',....\theta_n'$ is less than θ

Solution(cont'd)

-> Choose the magnitude of the unprimed vectors and calculate the resultant angle using the equation:

$$\tan(\theta) = (\Sigma v_i \sin\theta_i) / (\Sigma v_i \cos\theta_i)$$

-> We can always find primed vectors' angles using the following boundary conditions:

$\theta_i' > \theta_i$ and at least one of $\theta_i' < \theta$

Solution(cont'd)

-> Let θ_1' , θ_2' , ……,θ_x' < θ and θ_{x+1}',θ_{x+2}',….,θ_n' > θ , Pick a value R'_{max} such that

$$\tan\theta = \left(\left(\Sigma^{x}_{i=1} v_i' \sin\theta_i'\right) + R_{max}'\sin\theta_a'\right) / \left(\left(\Sigma^{x}_{i=1} v_i'\cos\theta_i'\right) + R_{max}'\cos\theta_a'\right)$$

-> Pick a Value of $R' < R_{max}'$ and calculate the angle θ'. It turns out to be less than θ.

-> Express **R'** (which makes angle θ_a with the horizontal axis) as a linear combinations of vectors $\mathbf{v}'_{x+1}, \mathbf{v}'_{x+2}, \ldots, \mathbf{v}'_n$.

‘

Examples

-> UC Berkeley Gender Bias

->Kidney Stone Treatment