Selection and the Roy Model – Part I

John Eric Humphries

April 22, 2014

An Island Economy

Consider an example you are all familiar with. Let there be an island economy where people either hunt parrots or fish. This may seem annoyingly simple, but a two-occupation economy serves as the building block to much bigger models.

We will assume no uncertainty in the number of fish or parrots an individual would catch and that all fish and all parrots are the same.

Also assume a competitive market for fish and parrots with:

- F_i is the number of fish caught by person i
- B_i is the number of parrots caught by person i (using B for "bird" to avoid using P, which could be confused for prices).
- p_F is the price of fish
- p_B is the price of birds
- (from here on out we will supress the i subscripts for simpler notation, but remember, skills are individual specific, while prices apply to the whole market.)

Thus, we have that an individual can choose between $W_F = p_F F$ or $W_B = p_B B$. Thus, the individual's wage will be:

$$W = max\{W_F, W_B\}$$

Who hunts and who fishes?

An individual will be indifferent to hunting versus fishing when:

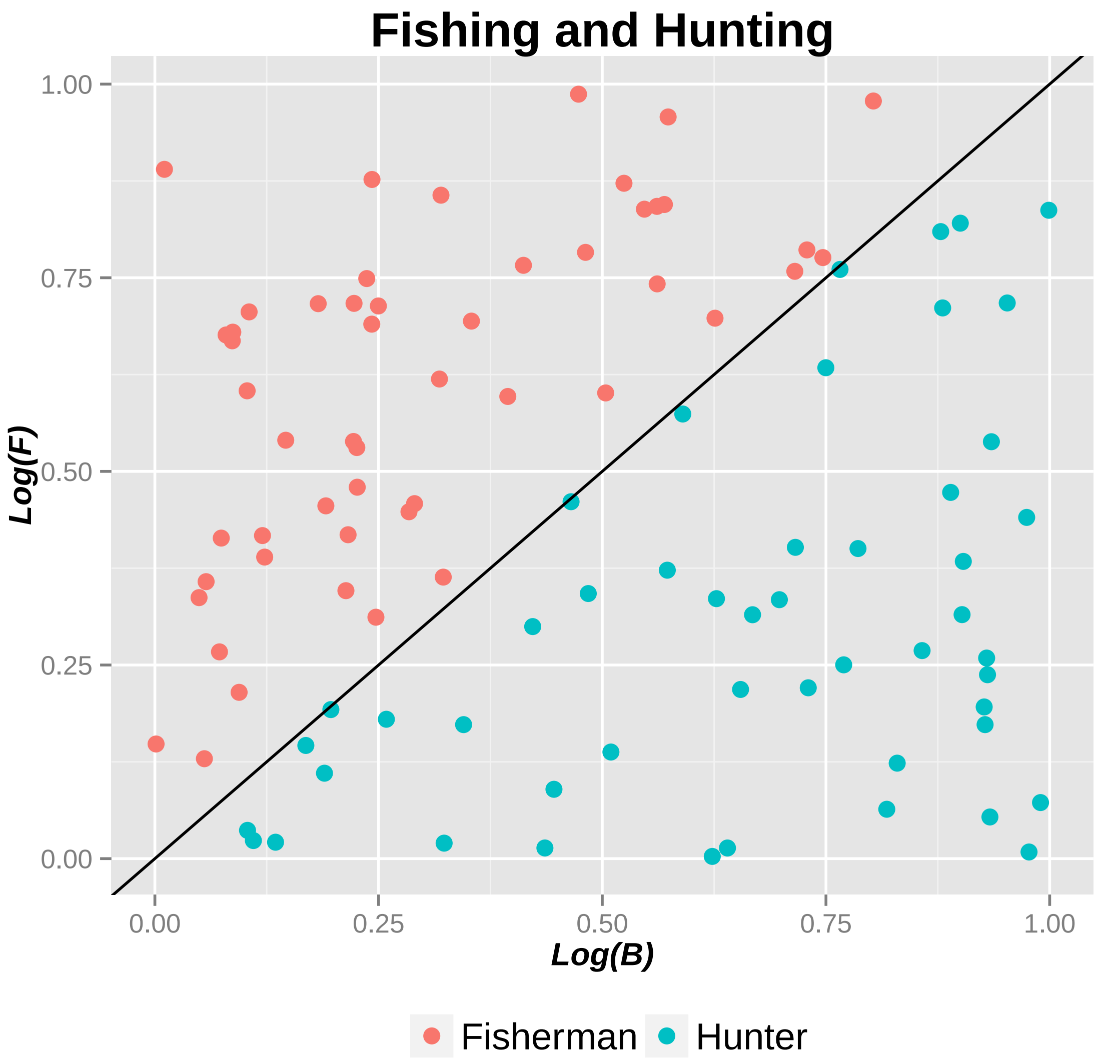
$$p_F F = p_B B$$

taking logs we can rewrite this as:

$$log(p_F) + log(F) = log(p_B) + log(B)$$

$$\Rightarrow log(F) = \underbrace{log(p_B) - log(p_F)}_{constant} + log(B)$$

Note that this equation is a line with a slope of 1 in the $log(B) \times log(F)$ space. If you are above this line you fish, and if you are below this line you hunt:

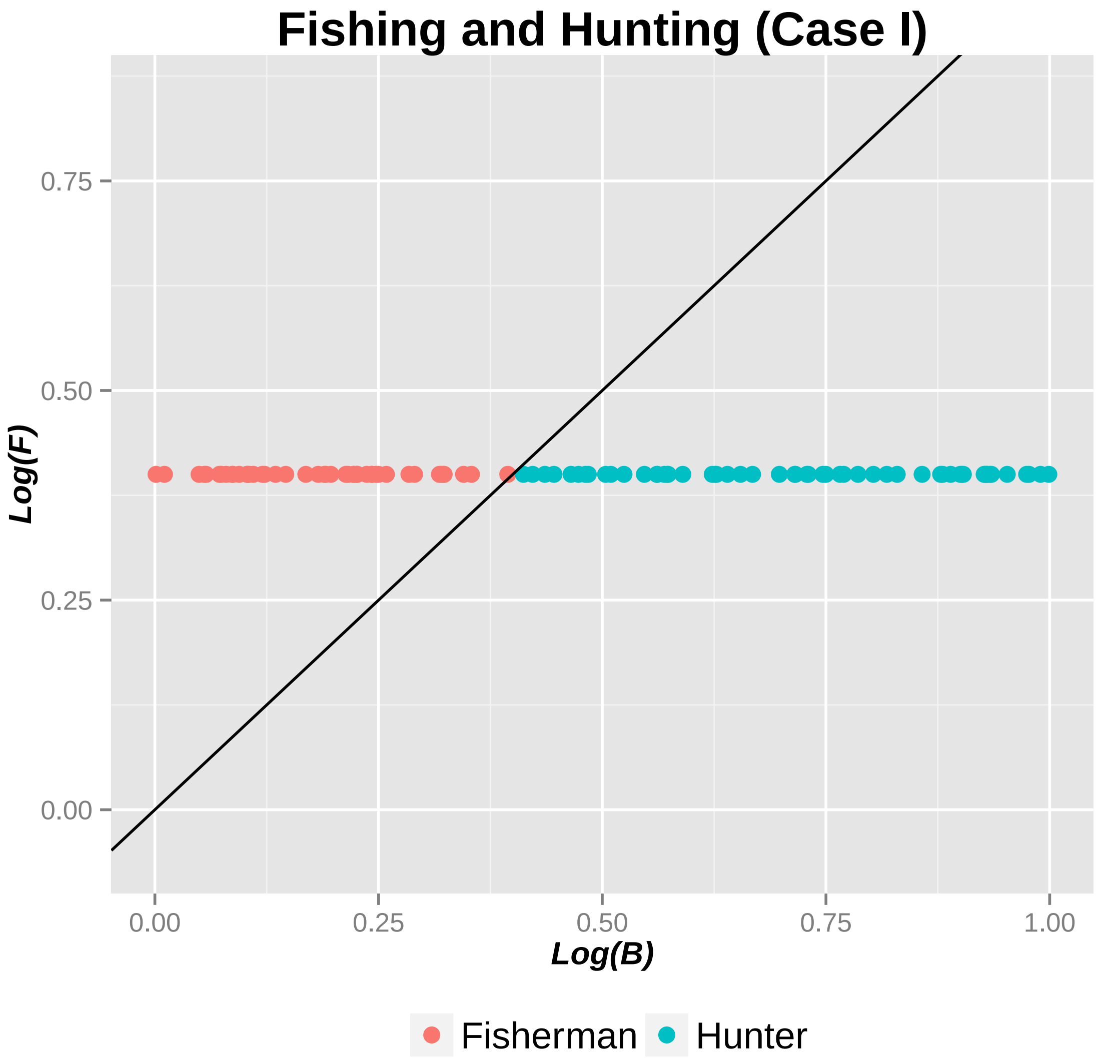


Do the best fishemen fish? Do the best hunters hunt? The answer will depend on the covariance matrix of skills. (1) Skill with larger log variance will have more sorting. (2) Demand does not matter!

Case 1: No variance in fishing

Consider the simple case where everyone who fishes catches $F*$ fish. Then, fishing receiveds $W* = p_F F*$. Returning to our sorting equation, we get that you will:

- Hunt if $B > \frac{W*}{p_B}$
- Fish if $B \leq \frac{W*}{p_B}$
- So the best hunters will hunt and all be better off than if they fished and everyone else who fished.



Case 2: Perfect correlation between skills

Let there be perfect correlation between skills given by

$$log(F) = \alpha + \beta log(B)$$

This implies that:

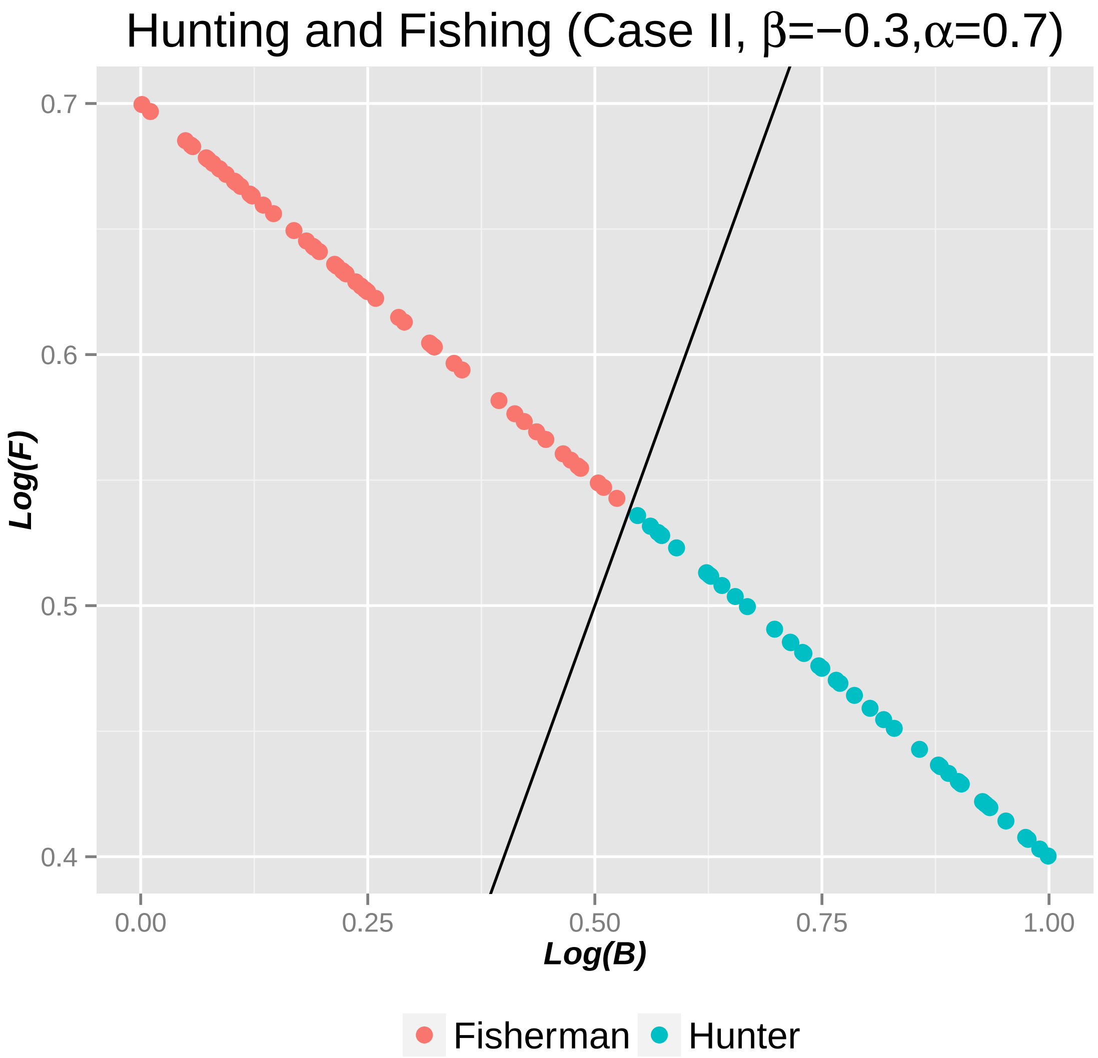
$$var(log(F)) = \beta^2 var(log(B))$$

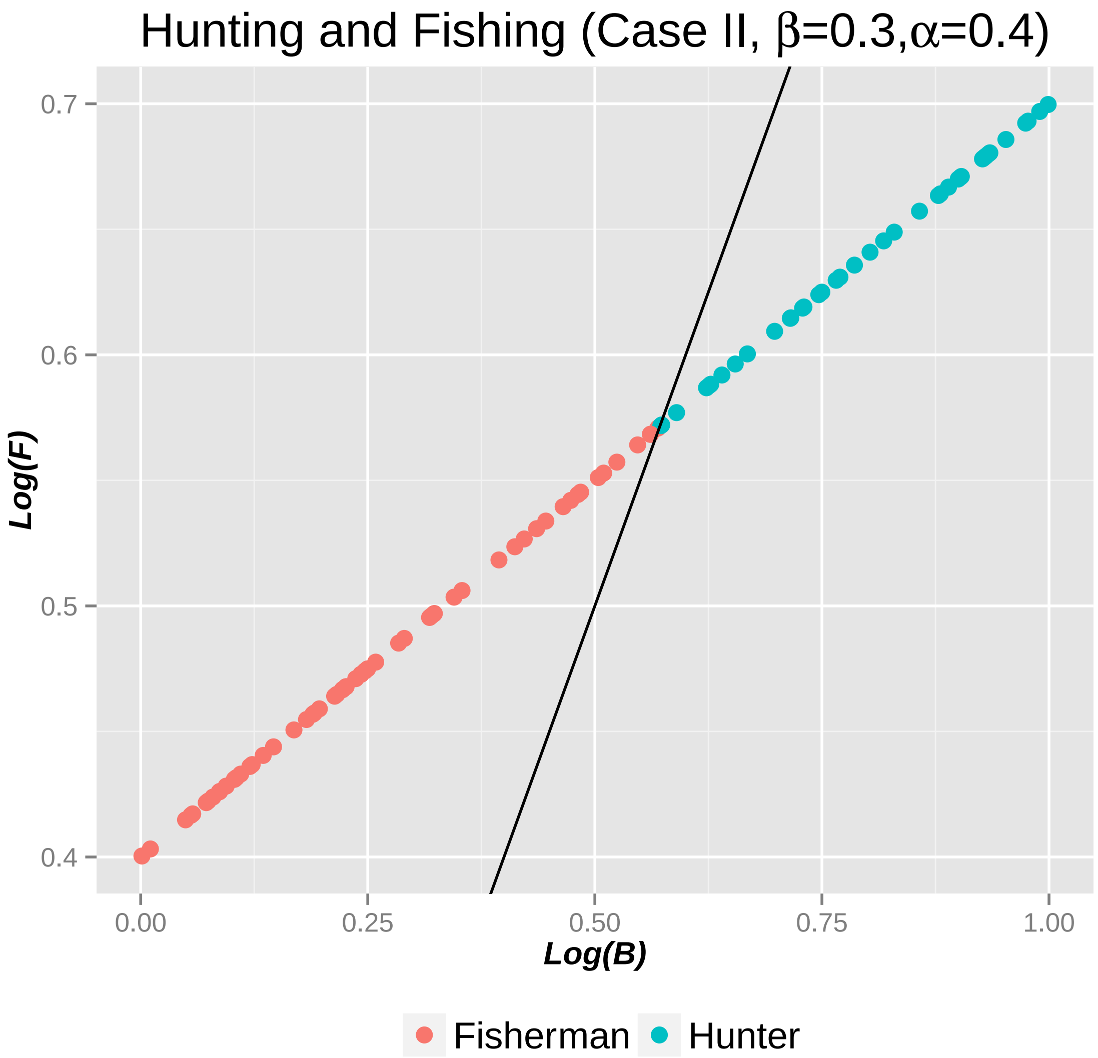
thus

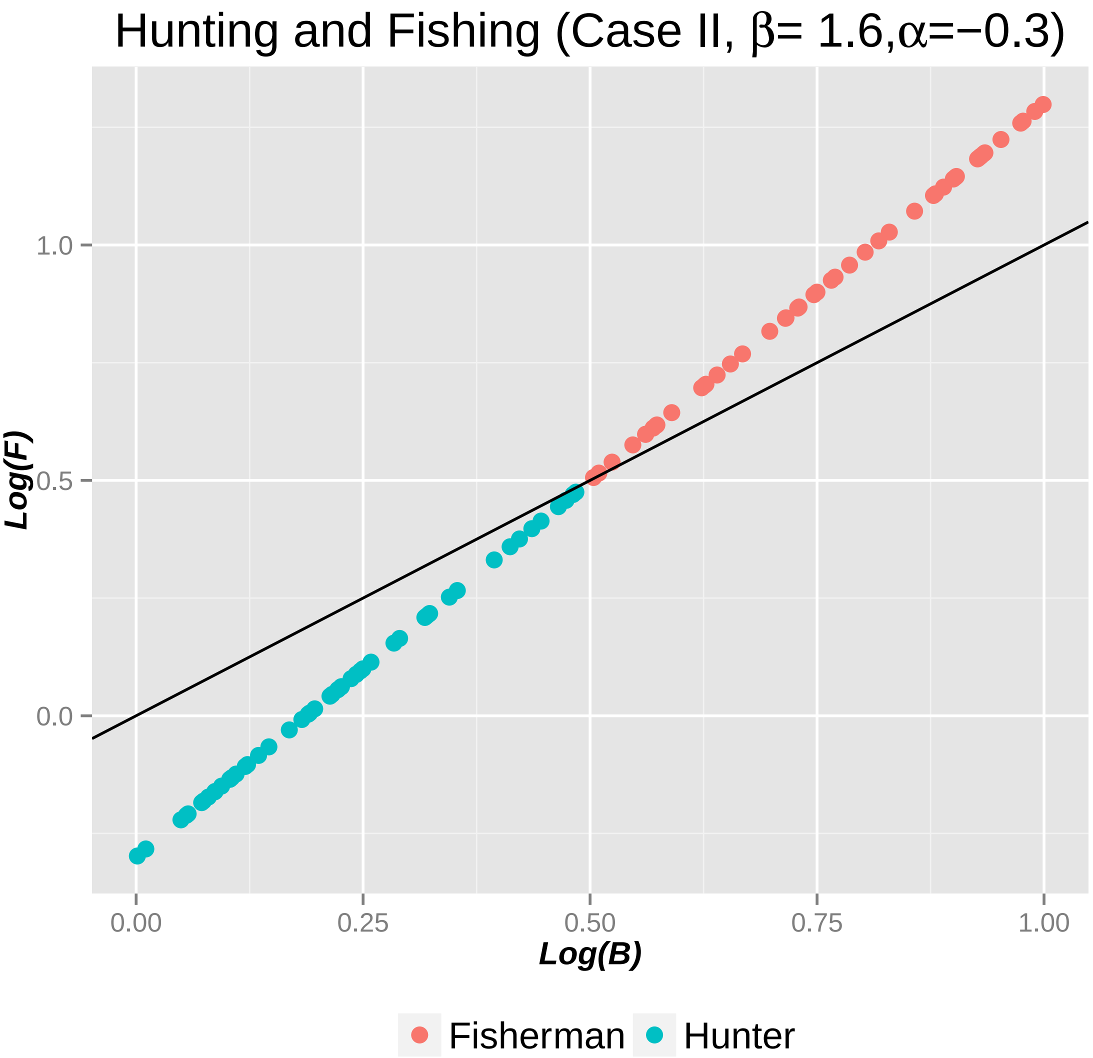
$$var(log(F)) < var(log(B)) \Rightarrow \beta < 1$$

We have three potential cases:

- $\beta \leq 0$
- $0 < \beta < 1$
- $beta \geq 1$


Hunting and Fishing (Case II, β=−0.3,α=0.7)
0.7
0.6
0.5
0.4
Log(F)
0.00
0.25
0.50
0.75
1.00
Log(B)
Fisherman
Hunter

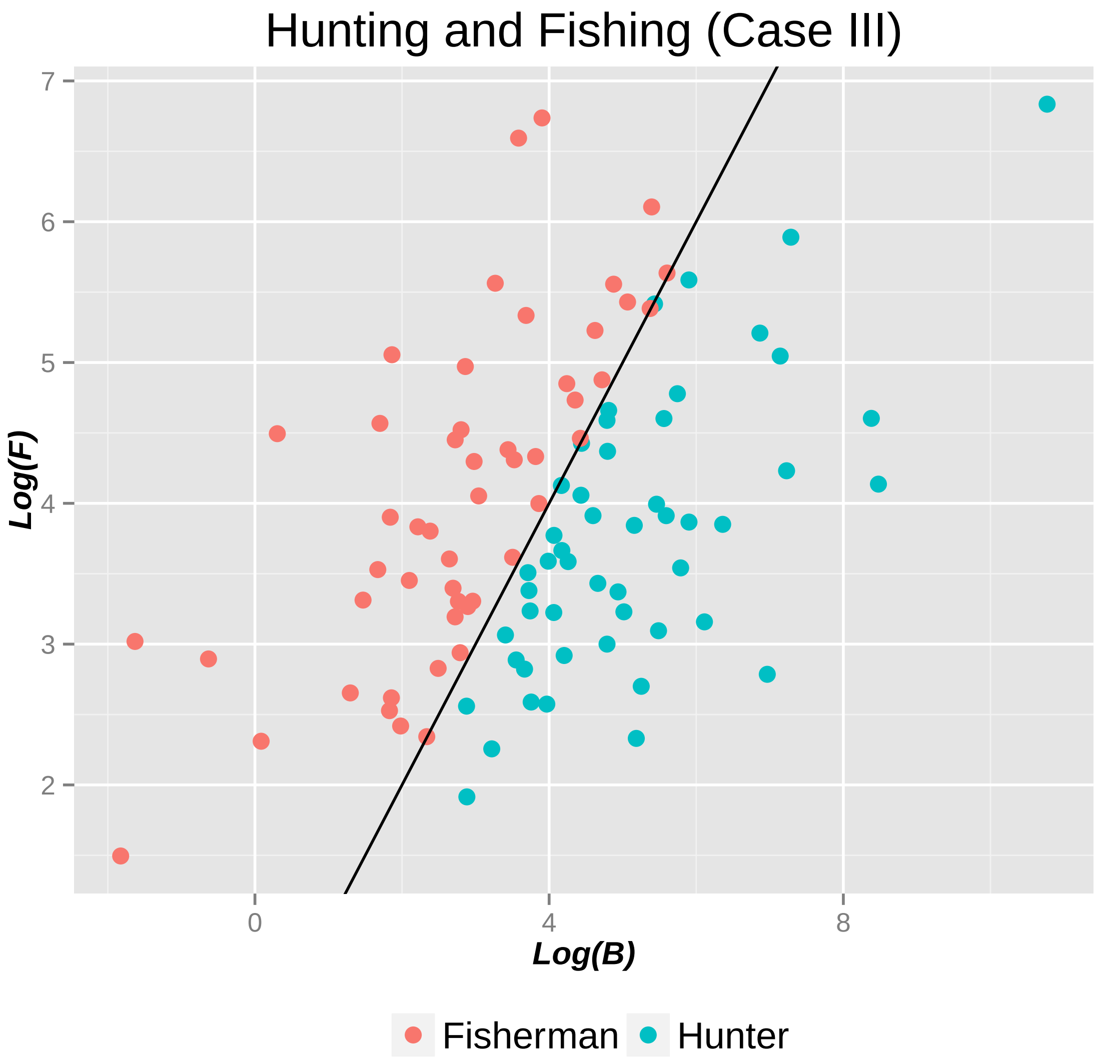

Hunting and Fishing (Case II, β=0.3,α=0.4)
0.7
0.6
0.5
0.4
0.00
0.25
0.50
0.75
1.00
Log(F)
Log(B)
Fisherman
Hunter

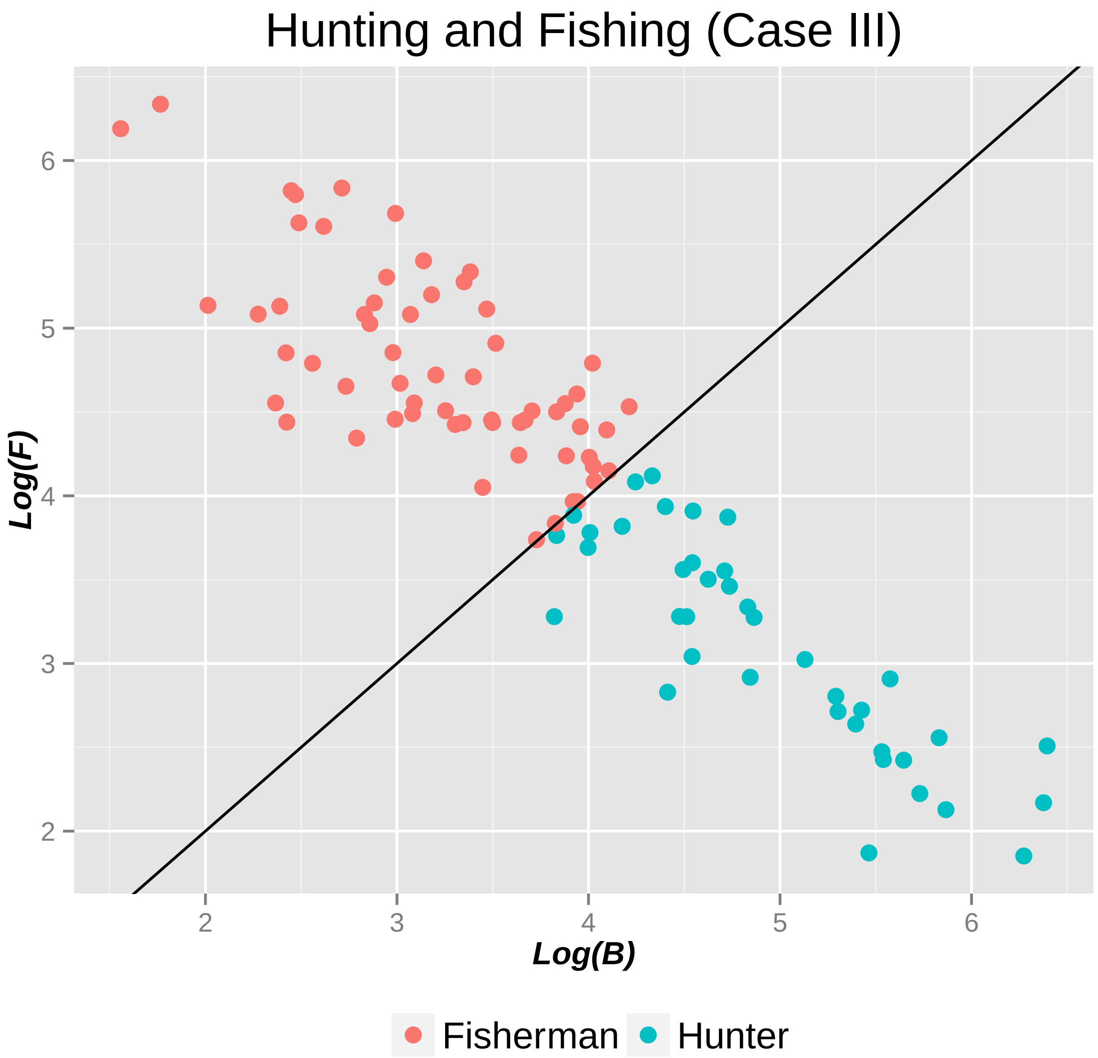


Case 3: Log Normal Random Variables

Now assume that

$$(log(B), log(F)) \sim N\left(\begin{bmatrix} \mu_B \\ \mu_F \end{bmatrix}, \begin{bmatrix} \sigma_{BB} & \sigma_{FB} \\ \sigma_{BF} & \sigma_{FF} \end{bmatrix}\right)$$


Hunting and Fishing (Case III)
Log(F)
Log(B)
7
6
5
4
3
2
0
4
8
Fisherman
Hunter



Adding more structure to the economy

Now we will consider a slightly more advanced (and realistic) model for our island economy with fishing and hunting production functions for individual i given by:

$$F_i = e^{X'_{i,f}\beta_f + \mu_{i,f}}$$

$$B_i = e^{X'_{i,b}\beta_b + \mu_{i,b}}$$

Note that the individual atributes (X) could vary by occupation. It could be that swimming ability is useful for fishing, while running ability is useful for hunting. The prices (β) can also vary by occupation. For example, intelligence may be more important for hunting.[1] We assume that as economists we observe the wage in the occupation they chose to work in and their observable characteristics X. We assume the individual knows $\mu_{i,f}$ and $\mu_{i,b}$ but that these are unobserved to the economist.

[1]Note that we can define $X = X_f \cup X_b$, whereh the price running ability is 0 in the fishing occupation and the price of swimming ability is 0 in the hunting occupation.

Pluggin in F_i and B_i into our wage equation and taking logs, we have:

$$log(W_{i,f}) = log(p_f) + X_i'\beta_f + \mu_{i,f}$$

$$log(W_{i,b}) = log(p_b) + X_i'\beta_b + \mu_{i,b}$$

A person will hunt if: $W_{i,b} > W_{i,f}$, which implies that the individual will choose to hunt if:

$$(log(p_b) - log(p_f)) + X_i'(\beta_b - \beta_f) + (\mu_{i,b} - \mu_{i,f}) > 0$$

- This is a "latent threshold" model.
- It is important to differentiate what is known by the economists from what is known (and acted on) by the agents.
- Can assume different X in different occupation, but can also assume that the coefficient
- If we have assumed $\mu_{i,b}$ and $\mu_{i,f}$ are both normal, then we know the distribution of $(\mu_{i,b} - \mu_{i,f})$ (which will also be normal).
- When we have assumed normality, this is a probit model, like the one from our first homework