Revised Yitzhaki

James J. Heckman and Rodrigo Pinto

Econ 312, Spring 2022

THE UNIVERSITY OF CHICAGO

Yitzhaki's Theorem

Let Y, X be r.v. such that $E(|Y|) < \infty$, $E(|X|) < \infty$.
If $E(Y|X)$ is differentiable, then:

$$
\begin{aligned}
Cov(Y,X) &= \int_{-\infty}^{\infty} \frac{\partial E(Y|X=x)}{\partial x} \cdot E((X-\mu_X)\mathbf{1}[X>x])dx \\
&= \int_{-\infty}^{\infty} \frac{\partial E(Y|X=x)}{\partial x} \cdot E(X-\mu_X|X>x)P(X>x)dx \\
&= \int_{-\infty}^{\infty} \frac{\partial E(Y|X=x)}{\partial x} \cdot E(X-\mu_X|X>x)(1-F_X(x))dx \\
&= \int_{-\infty}^{\infty} \frac{\partial E(Y|X=x)}{\partial x} \cdot \int_{x}^{\infty} (x-\mu_X)f_X(x)dx
\end{aligned}
$$

where $\mu_X = E(X)$, $F_X(x) = P(x \leq X)$ and $f_X(x)$ is the density of r.v. X.
A consequence is:

$$
Var(X) = \int_{-\infty}^{\infty} E(X-\mu_X|X>x)P(X>x)dx
$$

Applying Yitzhaki Weights to Standard OLS

$$Y = \beta_1 + \beta_2 X + \epsilon, \text{ such that } \epsilon \perp\!\!\!\perp X \text{ and } E(\epsilon) = 0$$

$$\Rightarrow E(Y|X = x) = \beta_2 x \text{ and } \beta_2 = \frac{Cov(X, Y)}{Var(X)}$$

What can we get from applying Yitzhaki weights?

$$Cov(X, Y) = \int \underbrace{\frac{\partial E(Y|X = x)}{\partial x}}_{\text{Constant Slope}} \cdot E(X - \mu_X | X > x) P(X > x) dx$$

$$= \beta_2 \int \cdot E(X - \mu_X | X > x) P(X > x) dx$$

$$= \frac{Cov(X, Y)}{Var(X)} \int \cdot E(X - \mu_X | X > x) P(X > x) dx$$

$$\Rightarrow Var(X) = \int E(X - \mu_X | X > x) P(X > x) dx$$

But this is a result we already know !

What about the non-linear case?

Consider a general case of a non-linear relation between X and Y:

$$Y = g(X) + \epsilon, \text{ such that } E(\epsilon|X) = 0$$
$$\Rightarrow E(Y|X = x) = g(x)$$

- The slope is **not** constant, but given by $\frac{\partial E(Y|X=x)}{\partial x} = g'(x)$
- **Question:** What is the interpretation of OLS estimator $Cov(X, Y)/Var(X)$ in terms of the slope $g'(x)$?

What does OLS **Really** Evaluates?

What does OLS Really Evaluates?

OLS evaluates a weighted average of the slope $\frac{\partial E(Y|X=x)}{\partial x} = g'(x)$

$$\frac{Cov(X,Y)}{Var(X)} = \int \underbrace{\frac{\partial E(Y|X=x)}{\partial x}}_{\text{Slope}} \cdot \underbrace{\frac{E(X-\mu_X|X>x)P(X>x)}{Var(X)}}_{\text{Positive Weights}} dx$$

$$= \int \underbrace{g'(x)}_{\text{Slope}} \cdot \underbrace{w(x)}_{\text{Weights}} dx$$

1. Weights sum to 1:

$$E(w(x)) = \frac{\int E(X-\mu_X|X>x)P(X>x)dx}{Var(X)} = \frac{Var(X)}{Var(X)} = 1$$

2. Weights are positive:

$$w(x) = E(X-\mu_X|X>x)P(X>x)$$

$$\Rightarrow w(x) < w(x') \text{ wherever } x < x'$$

$$\lim_{x\to-\infty} E(w(x)) = E(X-\mu_X|X>-\infty)P(X>-\infty) = E(X-\mu_X)\cdot 1 = 0$$

Thus $w(-\infty) = 0 < w(x)$ for all $x \in \mathbb{R}$

Additional Expressions of Covariance Yatracos (1998)

$$Cov(Y,X) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y(x-\mu_x) f_{X,Y}(x,y)dxdy$$

$$Cov(Y,X) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Big(F_{X,Y}(x,y) - F_X(x)F_Y(y)\Big)dxdy$$

$$Cov(Y,X) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} Cov(\mathbf{1}[X \geq x], \mathbf{1}[Y \geq y])dxdy,$$

where $F_Y(y) \equiv P(Y \leq y),\ F_X(x) \equiv P(X \leq x)$, $F_{X,Y}(x,y) \equiv P(Y \leq y, X \leq x)$ are CDFs and $f_{X,Y}(x,y)$ is the joint density of r.v. X, Y.

Expressions for Expectation and Variance

$$E(Y) = \int_{-\infty}^{+\infty} y\, f_Y(y) dy$$

$$E(Y) = \int_{0}^{+\infty} P(Y > y) dy - \int_{-\infty}^{0} P(Y \leq y) dy$$

$$Var(Y) = \int_{-\infty}^{\infty} \Big(F(y)(1 - F(y)) P(Y \leq y) \Big) \cdot \Big(E(Y|Y > y) - E(Y|Y \leq) \Big) dy$$

$$Var(Y) = 2 \int_{-\infty}^{\infty} F(y)(1 - F(y)) \cdot \Big(y - E(Y|Y < y) \Big) dy$$

$$Var(Y) = 2 \int \int_{-\infty < y < y' < \infty} F(y)(1 - F(y')) dy dy'$$

where $F_Y(y) \equiv P(Y \leq y),\ F_X(x) \equiv P(X \leq x)$,
$F_{X,Y}(x, y) \equiv P(Y \leq y, X \leq x)$ are CDFs and $f_{X,Y}(x, y)$ is the joint density of r.v. X, Y.