BIOS 667: Longitudinal Data Analysis
Double Expectation formulae

The *double expectation formulae* is a collection of strongly related formulae.

Suppose that the random variables (X, Y, Z) have a joint distribution, $A, B, A_1, \ldots$ are events, and use $I(A)$ for the indicator function of event A.

The fundamental double expectation formula is

$$\mathrm{E}[Y] = \mathrm{E}[\mathrm{E}[Y|X]].$$

Note that the three E's above are all different! They are as follows,

$$\mathrm{E}_Y[Y] = \mathrm{E}_X[\mathrm{E}_{Y|X}[Y|X]].$$

The first is over the marginal distribution of Y. The second is over the marginal distribution of X. The third is over the conditional distribution of Y given X. The subscripts are usually dropped to simplify the notation. But it is important to know that each of the three expectations is with respect to a different distribution.

Variance and covariance:

$$\mathrm{var}(Y) = \mathrm{E}[\mathrm{var}(Y|X)] + \mathrm{var}(\mathrm{E}[Y|X])$$

$$\mathrm{cov}(Y, X) = \mathrm{E}[\mathrm{cov}(Y, X|Z)] + \mathrm{cov}(\mathrm{E}[Y|Z], \mathrm{E}[X|Z])$$

Note that in the last formula, Z can be X giving

$$\mathrm{cov}(Y, X) = \mathrm{E}[\mathrm{cov}(Y, X|X)] + \mathrm{cov}(\mathrm{E}[Y|X], \mathrm{E}[X|X]) = \mathrm{E}[0] + \mathrm{cov}(\mathrm{E}[Y|X], X) = \mathrm{cov}(\mathrm{E}[Y|X], X)$$

If $Y = I(B)$, the above is

$$P(B) = \mathrm{E}[P[B|X]]$$

pdf's and cdf's:

$$f_Y(y) = \mathrm{E}[f_{Y|X}(y|X)], \qquad F_Y(y) = \mathrm{E}[F_{Y|X}(y|X)]$$

For any *partition* of the sample space into sets $A_1, \ldots, A_n$,

$$\mathrm{E}[Y] = \sum_{i=1}^{n} \mathrm{E}[Y|A_i]P(A_i).$$