Laws of Total Expectation & Variance (3 pages; 16/4/21)

[relating to discrete random variables]

(1) Conditional expectation, $E(X|Y)$ (for any random variables X & Y)

First of all consider $f(y) = E(X|Y = y) = \sum_x xP(X = x|Y = y)$

Then, considering y as a random variable, $E(X|Y) = f(Y)$.

Example 1 ("Randomly stopped sum"): $X = X_1 + X_2 + \cdots + X_N$, where X_i is independent of X_j for $i \neq j$, X_i is independent of N, and $E(X_i) = \mu$ and $Var(X_i) = \sigma^2$ for each i, and where N is equally likely to be each of the integers 1 to n. [Note though that the X_i don't need to have the same distribution.]

Here $Y = N$, and $E(X|N) = E(X_1 + X_2 + \cdots + X_N) = N\mu$

(2) $E\big(g(X)\big) = E_Y(E(g(X)|Y))$ for any function g

Proof

RHS = $E_Y\{\sum_x g(x)P(X = x|Y)\}$

$= \sum_y[\sum_x g(x)P(X = x|Y = y)]P(Y = y)$

$= \sum_y \sum_x g(x)P(X = x|Y = y)P(Y = y)$

$= \sum_x g(x) \sum_y P(X = x|Y = y)(Y = y)$

$= \sum_x g(x)P(X = x)$

$= E\big(g(X)\big)$

(3) Law of total expectation

For any random variables X & Y, $E(X) = E_Y\{E(X|Y)\}$ (*)

where $E_Y(E(X|Y)) = \sum_y E(X|Y = y)P(Y = y)$

$= \sum_y f(y)P(Y = y)$

[(*) applies to continuous random variables as well]

Proof: Set $g(X) = X$ in (2).

Example 1 (again):, $E(X) = E_N\{E(X|N)\} = E_N\{N\mu\} = \mu E_N\{N\}$

And $E_N\{N\}[\equiv E(N)] = \sum_{r=1}^{n} r\,.\ \frac{1}{n} = \frac{1}{n}\,.\frac{1}{2}\,n(n+1) = \frac{1}{2}(n+1)$

So $E(X) = \frac{\mu}{2}(n+1)$

(4) Law of total variance

For any random variables X & Y,

$Var(X) = E_Y(Var(X|Y)) + Var_Y(E(X|Y))$ (**)

[(**) applies to continuous random variables as well]

Proof

RHS $= E_Y\{E(X^2|Y) - [E(X|Y)]^2\}$

$+E_Y\{[E(X|Y)]^2\} - \{E_Y[E(X|Y)]\}^2$

$= E(X^2) - E_Y\{[E(X|Y)]^2\} + E_Y\{[E(X|Y)]^2\} - [E(X)]^2$

$= E(X^2) - [E(X)]^2 = Var(X)$

[See "PGF – Exercises" for a (longer) proof using probability generating functions.]

Example 1 (again): $Var(X) = E_N(Var(X|N)) + Var_N(E(X|N))$

Now, $Var(X|N) = Var(X_1 + X_2 + \cdots + X_N) = N\sigma^2$,

and so $E_N(Var(X|N) = E_N(N\sigma^2) = \sigma^2.\frac{1}{2}(n+1)$

And $Var_N\left(E(X|N)\right) = Var_N(N\mu) = \mu^2 Var_N(N)$,

and $Var_N(N) = E(N^2) - [E(N)]^2$

$= \{\sum_{r=1}^{n} r^2 \,.\frac{1}{n}\} - [\frac{1}{2}(n+1)]^2$

$= \frac{1}{n}\,.\frac{1}{6}n(n+1)(2n+1) - \frac{1}{4}(n+1)^2$

$= \frac{(n+1)}{12}[2(2n+1) - 3(n+1)]$

$= \frac{(n+1)(n-1)}{12}$

So $Var(X) = \frac{\sigma^2}{2}(n+1) + \frac{\mu^2(n+1)(n-1)}{12}$

Example 2: As Example 1, but N has a more general distribution.

$Var(X) = E_N(N\sigma^2) + \mu^2 Var_N(N) = \sigma^2 E(N) + \mu^2 Var(N)$

$[E(N) \equiv E_N(N)$ & $Var(N) \equiv Var_N(N)]$

Where $N \sim Po(\lambda)$, for example,

$E(X) = \lambda\mu$ and $Var(X) = \lambda(\sigma^2 + \mu^2)$.